WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5: C12P 21/06, C12N 5/00, 15/00 C07H 15/12, 17/00, C07K 3/00 C07K 13/00, 15/00, 17/00 A61K 35/14

(11) International Publication Number:

WO 92/10583

A1

(43) International Publication Date:

25 June 1992 (25.06.92)

(21) International Application Number:

PCT/US91/09422

(22) International Filing Date:

12 December 1991 (12.12.91)

(30) Priority data: 626,806 648,481

672,007

12 December 1990 (12.12.90) 30 January 1991 (30.01.91) 18 March 1991 (18.03.91)

(60) Parent Application or Grant

(63) Related by Continuation US Filed on

626,806 (CIP) 12 December 1990 (12.12.90)

(71) Applicants (for all designated States except US): ZYMOG-ENETICS, INC. [US/US]; 4225 Roosevelt Way N.E., Seattle, WA 98105 (US). THE BOARD OF REGENTS OF THE UNIVERSITY OF WASHINGTON [US/US]; Seattle, WA 98195 (US).

(75) Inventors/Applicants (for US only): MULVIHILL, Eileen, Ranae [US/US]; 4016 Francis Avenue North, Seattle, WA 98103 (US). HAGEN, Frederick, Stammer [US/ US]; 1315 Lexington Way East, Seattle, WA 98112 (US). HOUAMED, Khaled, M. [TN/US]; 730 Harvard Avenue, Seattle, WA 98102 (US). ALMERS, Wolfhard [DE/US]; 825 35th Avenue, Seattle, WA 98122 (US).

(74) Agents: PARMELEE, Steven, W. et al.; Townsend and Townsend, One Market Plaza, 2000 Steuart Tower, San Francisco, CA 94105 (US).

(81) Designated States: AT, AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI patent), CH, CH (European patent), CI (OAPI patent), CM (OAPI patent), CS, DE, DE (European patent), DK, DK (European patent), ES, ES (European patent), FI, FR (European patent), GA (OAPI patent), GB, GB (European patent), GN (OAPI patent), GR (Furopean patent), GN (OAPI ropean patent), GN (OAPI patent), GR (European patent), HU, IT (European patent), JP, KP, KR, LK, LU, LU (European patent), MC (European patent), MG, ML (OAPI patent), MN, MR (OAPI patent), MW, NL, NL (European patent), NO, PL, RO, SD, SE, SE (European patent), SN (OAPI patent), SU⁺,TD (OAPI patent), TG (OAPI patent), US.

Published

With international search report,

(54) Title: G PROTEIN-COUPLED GLUTAMATE RECEPTORS

(57) Abstract

Mammalian G protein-coupled glutamate receptors are identified, isolated and purified. The receptors have been cloned, sequenced and expressed by recombinant means. The receptors and antibodies thereto can be used to identify agonists and antagonists of G protein-coupled glutamate receptor mediated neuronal excitation and in methods of diagnosis.

+ DESIGNATIONS OF "SU"

Any designation of "SU" has effect in the Russian Federation. It is not yet known whether any such designation has effect in other States of the former Soviet Union.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

	• • •		•		•		•
	AT.	Austria		es ·	Spain	MG	Madagascar
		Australia		Fi	Finland	ML.	Mali
	AU			FR	France	MN	Mongolia ·
	22	Barbados	-		Gabon	MR	Mauritania
	BE	Belgium		GA	United Kingdom	MW	Malawi
	BF	Burkina Faso		GB ·		NL	Netherlands
	BG	Bulgaria	_	GN	Guinea	NO	Norway
	BJ	Benin	-	GR	Greece		Poland
	BR	Brazil	ŀ	HU	Hungary	PL	
	CA	Canada	ı	IT.	Italy	RO	Romania
	CF	Central African Republic	1	JP	Japan	SD	Sudan
	CC	Congo		KP	Democratic People's Republic	SB	Sweden
	CH	Switzerland			of Korea	sn	Senegal
•				KR .	Republic of Korea	รบ+	Soviet Union
	CI	Côte d'Iveire	·	Li	Lischtenstein	TD	Chad
	CM	Cameroon			Sri Lanka	TG	Togo
	CS	Czochoslovakia	_	LK		US	United States of America
٠.	DE	Germany	_	FD .	Laxembourg		
	DK	Donassek		MC	Monuco		

G PROTEIN-COUPLED GLUTAMATE RECEPTORS

10

15

20

25

30

35

Background of the Invention

The majority of nerve cell connections are chemical synapses. A neurotransmitter is released from the presynaptic terminal, typically in response to the arrival of an action potential in the neuron, and diffuses across the synaptic space to bind to membrane receptor proteins of the postsynaptic terminal. The binding of neurotransmitters to membrane receptors is coupled either to the generation of a permeability change in the postsynaptic cell or to metabolic changes.

Neurotransmitters produce different effects according to the type of receptor to which they bind. In general, those which produce effects that are rapid in onset and brief in duration bind to receptors that act as ligand-gated ion channels, where binding almost instantly causes an ion flow across the membrane of the postsynaptic cell. Those neurotransmitters which act more like local chemical mediators bind to receptors that are coupled to intracellular enzymes, thereby producing effects that are slower in onset and more prolonged. These neurotransmitters alter the concentration of intracellular second messengers in the postsynaptic cell.

Four second messenger systems have been linked to neurotransmitter or hormone receptors and have been studied for their roles in the control of neuronal excitability. They are the adenylate cyclase/cyclic AMP-d pend nt protein kinase system, guanylat cyclase and cGMP-d pendent protein kinase, the in sit l trisph sphate/diacyl glycer l-pr t in kinase C system,

10

15

20

25

30

35

and systems which are activated by calcium ions, such as th calcium/calmodulin-dependent prot in kinas system. Thus, binding of a transmitter to a receptor may activate, for example, adenylate cyclase, thereby increasing the intracellular concentration of cAMP. The CAMP activates protein kinases that phosphorylate proteins in the cells, which form ion channels, thereby altering the cells' electrical behavior. As with the ligand-gated ion channel transmitters, the effects can be either excitatory or inhibitory, and may affect the cell at many levels, including the pattern of gene expression. It is also believed that these chemical synapses, associated with second-messenger systems, may be involved in long-term changes that comprise the cellular basis of learning and memory.

The ligand-activated membrane receptors do not activate the second messenger systems directly, however, but via a membrane-bound protein, the GTP-binding protein (G protein), which binds GTP on the cytoplasmic surface of the cell membrane and thereby acts to couple adenylate cyclase to the membrane receptor. Neurotransmitter binding to the membrane receptor is believed to alter the conformation of the receptor protein to enable it to activate the G protein in the lipid bilayer, which then binds GTP at the cytoplasmic surface and produces a further change in the G protein to allow it to activate, e.g., an adenylate cyclase molecule to synthesize cAMP. When a ligand binds a receptor, an enzymatic cascade results as each receptor activates several molecules of G protein, which in turn activate more molecules of adenylate cyclase which convert an even larger number of ATPs to CAMP molecules, producing a substantial amplification from the initial event.

Glutamate, aspartate and their endogenous derivatives ar believed to be the predominant excitat ry n urotransmitters in the vertebrate central nervous system. (Krinjrvic, Phys. Rev. 54:418-540, 1974). Recently, glutamate has be n d scribed as playing a

10

15

20

25

30

35

major, widespr ad role in the control of neuroendocrin neurons, possibly controlling not only the neuroendocrine system but other hypothalamic regions as well. major subclasses of glutamate receptors have been described but their characterization has until recently been limited to pharmacological and electrophysiological functional analyses. See generally, Hollman et al., Nature 342:643-648 (1989) and Sommer et al., Science 249:1580-1585 (1990). Three of the receptors, the quisqualate (QA/AMPA), kainate (KA), and N-methyl-Daspartate (NMDA) receptors, are believed to be directly coupled to cation-specific ion channels and thus are classified as ligand-gated ionotropic receptors. fourth glutamate receptor binds some of the agonists of the ionotropic receptors (quisqualate and glutamate, but not AMPA) but has no shared antagonists, and is coupled to G protein. Thus, this receptor, referred to as the G protein-coupled glutamate receptor, or Gluck, is pharmacologically and functionally distinct from the other major glutamate receptors. This receptor has also been termed the metabotropic receptor.

Agonist binding to GlugR has been shown to result in the activation of a number of second messenger systems, depending on the system studied. One of the best characterized is the quisqualate activation of phospholipase C through a G protein-coupled interaction that leads to the stimulation of inositol phospholipid This activity has been studied in systems that measure the accumulation of radiolabeled inositol monophosphate in response to stimulation by glutamate. The systems typically use brain slices from regions such as the hippocampus, striatum, cerebral cortex and hypothalamus (Nicoletti, et al., Proc. Natl. Acad. Sci. <u>USA</u> 83:1931-1935 (1986), and Nicoletti, et al., <u>J.</u> Neurochem. 46:40-46 (1986)), neur nal cultures derived from embry nic mouse and rat cerebellum, corpus striatum and cerebral c rt x (Nicoletti et al., J. Neurosci. 6:1905-1911 (1986), Sladeczek et al., Nature 317:717-719

10

15

20

25.

30

35

Æ

(1985), Dumuis, et al., Nature 347:182-184 (1990), and Drejer et al., J. Neurosci. 7:2910-2916 (1987)) and rat brain synaptosomes (Recasens et al., Eur. J. Pharm. 141: 87-93 (1987), and Recasens et al., Neurochem. Int. 13:463-467 (1988)). A major disadvantage of each of these model systems is the difficulty in analyzing the pharmacological and functional activities of Gluck in an environment where other glutamate receptors and G protein-coupled receptors such as muscarinic and serotonin receptors are also present.

.The <u>Xenopus</u> oocyte system has been used to identify GlugR as a member of the family of G proteincoupled receptors. An endogenous inositol triphosphate second messenger-mediated pathway in the oocyte allows the detection of GlugR after injection of total rat brain mRNA, in that the oocyte responds to ligand via the oocyte G protein-coupled PLC-mediated activation of a chloride channel that can be detected as a delayed, oscillatory current by voltage-clamp recording (Houamed et al., Nature 310:318-321 (1984), Gunderson et al., Proc. Royal Soc. B221:127-143 (1984), Dascal et al., Mol. Brain Res. 1:301-309 (1986), Verdoorn et al., Science 238:1114-1116 (1987), Sugiyama et al., Nature 325:531-533 (1987), Hirono et al., Neuros. Res. 6:106-114 (1988), Verdoorn and Dingledine, Mol. Pharmacol. 34:298-307 (1988), and Sugiyama et al., Neuron 3:129-132 (1989)). Injection of region-specific brain mRNA and of size fractionated mRNA have suggested that GlugR may be a large mRNA (6-7 kb) and that it is enriched in the cerebellum (Fong et al., Synapse 2:657-665 (1988) and Horikoshi et al., Neurosci. Lett. 105:340-343 (1989)).

There remains considerable need in the art for isolated and purified Glu_GR, as well as systems capable of expressing Glu_GR separate from other neurotransmitter receptors. Further, it would be desirable to specifically identify the presence of Glu_GR in cells and tissues, thereby aveiding the time-consuming, complex and nonspecific functional electrophysiological and

15

20

25

30 -

35

pharmacol gical assays. It would also be desirable to screen and develop new agonists and/or antagonists specific for Glu_cR , but to date this has not been practical. Quite surprisingly, the present invention fulfills these and other related needs.

Summary of the Invention

The present invention provides isolated and substantially pure preparations of mammalian G protein-coupled glutamate receptors and fragments thereof. In preferred embodiments the receptors are coupled to a G protein in vertebrate cells, bind glutamate and quisqualate and thereby activate phospholipase C, and are capable of stimulating inositol phospholipid metabolism. Having provided such receptors in isolated and purified form, the invention also provides antibodies to the receptors, in the form of antisera and/or monoclonal antibodies.

In another aspect the invention provides the ability to produce the mammalian G protein-coupled glutamate receptors and polypeptides or fragments thereof by recombinant means, preferably in cultured eukaryotic cells. The expressed receptors or fragments may or may not have the biological activity of corresponding native receptors, and may or may not be coupled to a G protein in the cell used for expression. Accordingly, isolated and purified polynucleotides are described which code for the receptors and fragments thereof, where the polynucleotides may be in the form of DNA, such as cDNA, or RNA. Based on these sequences probes may be used to hybridize and identify these and related genes which encode mammalian G protein-coupled glutamate receptors. The probes may be full length cDNA or as small as from 14 to 25 nucleotide, more often though from about 40 to ab ut 50 r m re nucle tides.

In related emb diments the inventi n c ncerns

DNA c nstructs which c mprise a transcripti nal pr m t r,

a DNA sequenc which nc des the r c pt r r fragm nt,

15 .

20

25

30

and a transcriptional terminator, each operably linked for expression of the r cept r. For expression the construct may also contain at least one signal sequence. The constructs are preferably used to transform or transfect eukaryotic cells, more preferably mammalian cells which do not express endogenous G protein-coupled glutamate receptors. When bound by an appropriate ligand such as glutamate or quisqualate, the receptor may activate phospholipase C in the host cell via coupling to G protein. Further, for large scale production the expressed receptor may also be isolated from the cells by, for example, immunoaffinity purification.

Cells which express the G protein-coupled glutamate receptors may also be used to identify compounds which can alter the receptor-mediated metabolism of a eukaryotic cell. Compounds may be screened for binding to the receptor, and/or for effecting a change in receptor-mediated metabolism in the host cell. Agonists and/or antagonists of the G protein-coupled glutamate receptors may also be screened in cell-free systems using purified receptors or binding fragments thereof for the effect on ligand-receptor interaction, or using reconstituted systems such as micelles which also provide the ability to assess metabolic changes.

In yet other embodiments the invention relates to methods for diagnosis, where the presence of a mammalian G protein-coupled glutamate receptor in a biological sample may be determined. For example, a monospecific antibody which specifically binds a G protein-coupled glutamate receptor is incubated with the sample under conditions conducive to immune complex formation, which complexes are then detected, typically by means of a label such as an enzyme, fluorophore, radionuclide, chemiluminescer, particle, or a second labeled antibody. Thus, means are provided for immunohistochemical staining of tissues, including brain tissues, for the subject receptors.

Brief Description of the Figures

5.

Figure 1 illustrates the construction of plasmid pVEGT, where Fig. 1A shows the construction of pVEG, Fig. 1B shows the construction of pVEG' and Fig. 1C shows pVEGT'. Symbols used are T7 pro, the T7 promoter; T1 and T2, synthetic and native T7 terminators, respectively; M13, M13 intergenic region; the parentheses indicate a restriction site destroyed in vector construction; and pA is the Aspergillus niger polyadenylate sequence.

15

20

25

30

35

10

Figure 2 illustrates representative responses from voltage-clamp assays of oocytes injected with RNA from positive pools.

Figure 3 illustrates a partial restriction map of clone 45-A.

Figure 4 illustrates the cloning of the receptor cDNA present in clone 45-A into Zem228R.

Figure 5 illustrates the DNA sequence and deduced amino acid sequence of clone 45-A (corresponding to Sequence ID Nos. 1 and 2). Numbers below the line refer to amino acid sequence, numbers above the line refer to nucleotide number. Putative transmembrane domains have been overlined, and putative N-linked glycosylation sites are indicated by closed circles.

Figure 6 illustrates a representative dose response curve for varying concentrations of L-glutamic acid. Error bars, where larger than the symbols, represent SEM.

Figure 7 illustrates the DNA sequence and deduced amino acid sequence of a subtype 1b glutamate receptor clone (Sequence ID Nos. 16 and 17). Numbers below the lin refer to amin acid sequence. Numbers ab ve the lin refer to nucleotid sequence.

Figur 8 illustrates the DNA sequ nce and deduced amino acid sequence of a subtype 2a glutamate

SUBSTITUTE SHEET

20

25

30

35

receptor cl ne (Sequence ID Nos. 18 and 19). Numb rs below th line refer to amino acid sequence. Numbers above the line refer to nucleotide sequence.

Figure 9 illustrates the DNA sequence of a partial subtype 2b glutamate receptor clone (Sequence ID No. 20). Numbers refer to the nucleotide sequence.

Description of the Specific Embodiments

Glu_GR is a family of G protein-coupled membrane receptors for the neurotransmitter glutamate. As glutamate has been described as having a major role in the control of neurons, particularly neuroendocrine neurons, Glu_GR may play a critical role in effectuating such control. Consequently, the development of agonists and antagonists of the Glu_GR-ligand interaction and Glu_GR-mediated metabolism is of great interest.

The present invention presents the means to identify agonists and antagonists of the GlugR-ligand interaction by providing isolated GlugR. The term "GlugR" refers to any protein either derived from a naturally occurring GlugR, or which shares significant structural and functional characteristics peculiar to a naturally occurring GluGR. Such a receptor may result when regions of a naturally occurring receptor are deleted or replaced in such a manner as to yield a protein having a similar function. Homologous sequences, allelic variations, and natural mutants; induced point, deletion, and insertion mutants; alternatively expressed variants; proteins encoded by DNA which hybridize under high or low stringency conditions to nucleic acids which encode naturally occurring GlugR-encoding nucleic acids; proteins retriev d from naturally occurring mat rials; and clos ly relat d proteins retrieved by antisera directed against GluGR proteins ar also included.

10

15

20

25

30

35

analog, or chimeric Glu_cR as gen rally described in U.S. Pat. No. 4,859,609, incorporated by reference herein. The molecule may be chemically synthesized or may occur in nature. Ligands may be grouped into agonists and antagonists. Agonists are those molecules whose binding to a receptor induces the response pathway within a cell. Antagonists are those molecules whose binding to a receptor blocks the response pathway within a cell.

By "isolated" Glu_cR is meant to refer to a Glu_cR which is in other than its native environment such as a neuron, including, for example, substantially pure Glu_cR as defined hereinbelow. More generally, isolated is meant to include a Glu_cR as a heterologous component of a cell or other system. For example, a Glu_cR may be expressed by a cell transfected with a DNA construct which encodes the Glu_cR, separated from the cell and added to micelles which contain other selected receptors. In another example described below, a Glu_cR is expressed by a cell which has been co-transfected with a gene encoding muscarinic receptor. Thus, in this context, the environment of the isolated Glu_cR is not as it occurs in its native state, particularly when it is present in a system as an exogenous component.

The invention provides cloned Glu_GR coding sequences which are capable of expressing Glu_GR proteins. Complementary DNA encoding Glu_GR may be obtained by constructing a cDNA library from mRNA from, for example, brain tissue. The library may be screened by transcribing the library and injecting the resulting mRNA into occytes and detecting, by functional assays, those occytes which express the Glu_GR. Alternatively, the clones may be screened with a complementary labeled oligonucleotide probe.

The present invention relates to successfully isolating a cDNA encoding a Glu_cR. Functional cloning of Glu_cR was acc mplish d by substantial modifications and improvem nts to a number of cDNA cloning and molecular biology techniques. Initially, an enriched source of

10

15

20

25

30

35

Glu R mRNA prepared by sucr se gradi nt centrifugation of >4kb length rat cerebellum poly(A)+ mRNA was used as template for cDNA synthesis. Further, a cDNA cloning vector that was employed included a poly(A) tail, thereby increasing by 40-fold the translational efficiency of the transcription product of the cDNA insert and a polylinker site to allow the directional cloning of the cDNA into the vector between the promoter and the poly(A) tail. Vector construction for directional cloning is described in co-pending U.S.S.N. 07/320,191, incorporated herein by The cDNA cloning vector also was used with two transcriptional terminators, in tandem, following the poly(A) sequences, efficiently generating a unit length transcript product without non-coding plasmid or viral sequences, and without requiring a restriction endonuclease to linearize the DNA template (a standard practice that will often prevent functional cloning strategies from working due to the presence of the endonuclease site within the coding region of the cDNA). The cDNA synthesis strategy maximized insert size and recreation of the 5' ends of the cDNA's, without introduction of homopolymer tails. cDNA inserts were size-selected to be greater than 4 kb in length before insertion into the vector. A library of 106 cDNA inserts in pools of 100,000 was replica plated to reduce the number of amplification steps in the fractionation of sequentially smaller pools. Moreover, m1 muscarinic cDNA (another G protein-coupled receptor coupled to phosphoinositol metabolism) template was included in transcription reactions of the subfractionated pools so that before injection the in vitro transcripts from each pool could be assayed by Northern analysis to assess relative quantity and quality of the mRNA, and by voltage-clamp of oocytes as an internal positive control for ach oocyte not r sponding to quisqualate or The inclusion of a dilution of SEAP-VEGT (a secreted form of alkaline phosphatas) template in transcriptions was also employed so that occytes selected

15

20

25

30

35

for voltage-clamp analysis w r thos synthesizing higher levels of the co-injected Glu_GR mRNA. And further, low noise electrical recording techniques were used to monitor the small signals initially generated from rare transcripts.

The above-described methods were used to isolate a cDNA clone encoding a Glu_GR designated "subtype la." Oligonucleotide probes based on the sequence of the subtype la clone were used to probe additional brain and cerebellum cDNA libraries. These libraries yielded clones encoding additional subtypes, which were designated lb, 2a and 2b.

With the Gluck and cDNA clones thereof provided herein, nucleotide and amino acid sequences may be determined by conventional means, such as by dideoxy sequencing. See generally, Sambrook et al., Molecular Cloning, A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989, incorporated by reference herein. Genomic or cDNA sequences encoding GlugR and homologous receptors of this family may be obtained from libraries prepared from other mammalian species according to well known procedures. For instance, using oligonucleotide probes from rodent Gluck, such as whole length cDNA or shorter probes of at least about fourteen nucleotides to twenty-five or more nucleotides in length; often as many as 40 to 50 nucleotides, DNA sequences encoding GlugR of other mammalian species, such as lagomorph, avian, bovine, porcine, murine, etc. may be obtained. If partial clones are obtained, it is necessary to join them in proper reading frame to produce a full length clone, using such techniques as endonuclease cleavage, ligation and loopout mutagenesis.

A DNA sequence encoding Glu_cR is inserted into a suitable expression vector, which in turn is used to transfect eukaryotic cells. Expression v ctors f r use in carrying out the pr s nt invention will comprise a

promoter capable of directing the transcription of a cloned DNA and a transcriptional terminator.

To direct proteins of the present invention for transport to the plasma membrane, at least one signal sequence is operably linked to the DNA sequence of interest. The signal sequence may be derived from the GlugR coding sequence, from other signal sequences described in the art, or synthesized de novo.

Host cells for use in practicing the present invention include mammalian, avian, plant, insect and fungal cells, but preferably mammalian cells. Fungal cells, including species of yeast (e.g., Saccharomyces spp., particularly S. cerevisiae, Schizosaccharomyces spp.) or filamentous fungi (e.g., Aspergillus spp., Neurospora spp.) may be used as host cells within the present invention. Suitable yeast vectors for use in the present invention include YRp7 (Struhl et al., Proc. Natl. Acad. Sci. USA 76: 1035-1039, 1978), YEp13 (Broach et al., Gene 8: 121-133, 1979), POT vectors (Kawasaki et al, U.S. Patent No. 4,931,373, which is incorporated by reference herein), pJDB249 and pJDB219 (Beggs, Nature 275:104-108, 1978) and derivatives thereof. Such vectors will generally include a selectable marker, which may be one of any number of genes that exhibit a dominant phenotype for which a phenotypic assay exists to enable transformants to be selected. Preferred selectable markers are those that complement host cell auxotrophy, provide antibiotic resistance or enable a cell to utilize specific carbon sources, and include LEU2 (Broach et al., ibid.), <u>URA3</u> (Botstein et al., <u>Gene</u> 8: 17, 1979), <u>HIS3</u> (Struhl et al., ibid.) or POT1 (Kawasaki et al., ibid.). Another suitable selectable marker is the CAT gene, which confers chloramphenicol resistance on yeast cells.

Additional vectors, promoters and terminators for use in xpressing the receptor of th invention in yeast are well known in the art and are reviewed by, for example, Emr, Meth. Enzymol. 185:231-279, (1990), incorporated herein by reference. The receptors of the

10

5

15

20

25

30

35

15

20

25

30

35

invention may be expressed in <u>Aspergillus</u> spp. (McKnight and Upshall, describ d in U.S. Pat nt 4,935,349, which is incorporated herein by reference). Useful promoters include those derived from <u>Aspergillus</u> nidulans glycolytic genes, such as the <u>ADH3</u> promoter (McKnight et al., <u>EMBO J.</u> 4:2093-2099, 1985) and the <u>tpiA</u> promoter. An example of a suitable terminator is the <u>ADH3</u> terminator (McKnight et al., ibid.). Techniques for transforming fungi are well known in the literature, and have been described, for instance by Beggs (ibid.), Hinnen et al. (<u>Proc. Natl. Acad. Sci. USA</u> 75:1929-1933, 1978), Yelton et al. (<u>Proc. Natl. Acad. Sci. USA</u> 81:1740-1747, 1984), and Russell (<u>Nature</u> 301:167-169, 1983) each of which are incorporated herein by reference.

A variety of higher eukaryotic cells may serve as host cells for expression of the Gluck, although not all cell lines will be capable of functional coupling of the receptor to the cell's second messenger systems. Cultured mammalian cells, such as BHK, CHO, Y1 (Shapiro et al., TIPS Suppl. 43-46 (1989)), NG108-15 (Dawson et al., Neuroscience Approached Through Cell Culture, Vol. 2, pages 89-114 (1989)), N1E-115 (Liles et al., J. Biol. Chem. 261:5307-5313 (1986)), PC 12 and COS-1 (ATCC CRL 1650) are preferred. Preferred BHK cell lines are the tk ts13 BHK cell line (Waechter and Baserga, Proc. Natl. Acad. Sci. USA 79:1106-1110 (1982)) and the BHK 570 cell line (deposited with the American Type Culture Collection, 12301 Parklawn Dr., Rockville, MD. under accession number CRL 10314). A tk BHK cell line is available from the ATCC under accession number CRL 1632.

Mammalian expression vectors for use in carrying out the present invention will include a promoter capable of directing the transcription of a cloned gene or cDNA. Preferred promoters include viral promoters and c llular promoters. Viral promoters include the imm diated array cytomegalovirus promoter (Boshart et al., Cell 41: 521-530, 1985) and the SV40 promoter (Subramani et al., Mol. Cell. Biol. 1: 854-864,

15

20

25

30

35

metallothion in-1 promoter (Palmiter et al., U.S. Pat nt No. 4,579,821), a mouse V_K promoter (Bergman et al., Proc. Natl. Acad. Sci. USA 81: 7041-7045, 1983; Grant et al., Nuc. Acids Res. 15: 5496, 1987) and a mouse V_H promoter (Loh et al., Cell 33: 85-93, 1983). A particularly preferred promoter is the major late promoter from Adenovirus 2 (Kaufman and Sharp, Mol. Cell. Biol. 2: 1304-13199, 1982). Such expression vectors may also contain a set of RNA splice sites located downstream from the promoter and upstream from the DNA sequence encoding the peptide or protein of interest. Preferred RNA splice sites may be obtained from adenovirus and/or immunoglobulin genes.

Also contained in the expression vectors is a polyadenylation signal located downstream of the coding sequence of interest. Polyadenylation signals include the early or late polyadenylation signals from SV40 (Kaufman and Sharp, ibid.), the polyadenylation signal from the Adenovirus 5 E1B region and the human growth hormone gene terminator (DeNoto et al., Nuc. Acids Res. 9: 3719-3730, 1981). The expression vectors may include a noncoding viral leader sequence, such as the Adenovirus 2 tripartite leader, located between the promoter and the RNA splice sites. Preferred vectors may also include enhancer sequences, such as the SV40 enhancer and the mouse μ enhancer (Gillies, Cell 33: 717-728, 1983). Expression vectors may also include sequences encoding the adenovirus VA RNAs.

Cloned DNA sequences may be introduced into cultured mammalian cells by, for example, calcium phosphate-mediated transfection (Wigler et al., Cell 14: 725, 1978; Corsaro and Pearson, Somatic Cell Genetics 7: 603, 1981; Graham and Van der Eb, Virology 52: 456, 1973.) Other techniques for introducing clon d DNA s quences into mammalian cells, such as electroporation (Neumann et al., EMBO J. 1: 841-845, 1982), may also be us d. In ord r to identify cells that have integrated

WO 92/10583 PCT/US91/09422

15

the cloned DNA, a sel ctable mark r is generally introduced into the cells along with the gene or cDNA of interest. Preferred selectable markers for use in cultured mammalian cells include genes that confer resistance to drugs, such as neomycin, hygromycin, and methotrexate. The selectable marker may be an amplifiable selectable marker. Preferred amplifiable selectable markers are the DAFR gene and the neomycin resistance gene. Selectable markers are reviewed by Thilly (Mammalian Cell Technology, Butterworth Publishers, Stoneham, MA, which is incorporated herein by reference). The choice of selectable markers is well within the level of ordinary skill in the art.

10

15

·20

25

30

35

Selectable markers may be introduced into the cell on a separate plasmid at the same time as the gene of interest, or they may be introduced on the same plasmid. If on the same plasmid, the selectable marker and the gene of interest may be under the control of different promoters or the same promoter, the latter arrangement producing a dicistronic message. Constructs of this type are known in the art (for example, Levinson and Simonsen, U.S. Patent No. 4,713,339). It may also be advantageous to add additional DNA, known as "carrier DNA" to the mixture which is introduced into the cells.

for a period of time, typically 1-2 days, to begin expressing the DNA sequence(s) of interest. Drug selection is then applied to select for growth of cells that are expressing the selectable marker in a stable fashion. Transfected cells may also be selected in the presence of antagonist to inhibit the activity of the receptor. Suitable antagonists in this context include D, L, 2-amino-3-phosphonopropionate. For cells that have been transfected with an amplifiable selectable marker the drug conc ntration may be incr as d in a st pwise manner to sel ct for incr ased copy number of the cloned sequences, thereby increasing xpression levels.

15

20.

25

30

35

Promoters, terminators and methods suitable for introducing expr ssion vectors encoding recombinant Glu_GR into plant, avian and insect cells are known in the art. The use of baculoviruses, for example, as vectors for expressing heterologous DNA sequences in insect cells has been reviewed by Atkinson et al. (Pestic. Sci. 28: 215-224,1990). The use of Agrobacterium rhizogenes as vectors for expressing genes in plant cells has been reviewed by Sinkar et al. (J. Biosci. (Banglaore) 11: 47-58, 1987).

Host cells containing DNA constructs of the present invention are then cultured to produce recombinant Gluck. The cells are cultured according to accepted methods in a culture medium containing nutrients required for growth of mammalian or other host cells. A variety of suitable media are known in the art and generally include a carbon source, a nitrogen source, essential amino acids, vitamins, minerals and growth factors. The growth medium will generally select for cells containing the DNA construct by, for example, drug selection or deficiency in an essential nutrient which is complemented by the selectable marker on the DNA construct.

Transfected cells expressing a cloned Gluck can be detected by several methods. By transfecting cells with an expression vector containing expression units for both the Gluck and a reporter gene (e.g. luciferase), the activity of the reporter gene provides an indicator of expression of the cotransfected Gluck clone. By including one or more cyclic AMP response elements (CRE) in the reporter gene expression unit, clones encoding receptors coupled to either the stimulation or inhibition of the second messenger adenylate cyclase can be identified by a change in reporter gene expression in response to added DNA constructs comprising a linked CRE and ligand. r porter gene ar known in the art. See, for example, Mellon et al., Proc. Natl. Acad. Sci. USA 86: 4887-4891 (1989), incorporated herein by reference. Cell lines

15

20

25

30

35

expr ssing functional recept rs can also be detected by el ctrophysiological measurements of agonist-induced channel activity. Receptor activity can also be assayed by measuring cytosolic free calcium concentrations in transfected cells. See, for example, Thastrup et al., Proc. Natl. Acad. Sci. USA 87: 2466-2470 (1990) and Picard et al., Science 247: 327-329 (1990), which are incorporated herein by reference. A preferred method for measuring cytosolic free calcium is by scanning cells with a fluorescent microscope coupled to a video camera. The cells are injected with a fluorescent Ca^{2*} indicator (e.g. Fluo-3 or Fura-2, Molecular Probes, Inc., Eugene, OR) and exposed to ligand.

The Gluck produced according to the present invention may be purified from the recombinant expression systems or other sources using purification protocols that employ techniques generally available to those skilled in the art. The most convenient sources for obtaining large quantities of Gluck are cells which express the recombinant receptor. However, other sources, such as tissues, particularly brain tissues of the cerebellum which contain Gluck, may also be employed.

Purification may be achieved by conventional chemical purification means, such as liquid chromatography, lectin affinity chromatography, gradient centrifugation, and gel electrophoresis, among others. Methods of protein purification are known in the art (see generally, Scopes, R., Protein Purification, Springer-Verlag, NY (1982), which is incorporated herein by reference) and may be applied to the purification of the Glu R and particularly the recombinantly produced Glu_R described herein. In a preferred embodiment immunoaffinity chromatography is employed using antibodies directed against GluR as herein described. another method of purification, a r combinant g ne encoding GlugR or portions ther of can be modified at the amino terminus, just behind a signal peptide, with a sequence coding for a small hydrophilic peptid , such as

15

20

25

described in U.S. Patent Nos. 4,703,004 and 4,782,137, incorporated herein by referenc. Specific antibodies for the peptide facilitate rapid purification of Glu_GR , and the short peptide can then be removed with enterokinase.

Thus, as discussed above, the present invention provides Glu_GR isolated from its natural cellular environment, substantially free of other G protein-coupled glutamate receptors. Purified Glu_GR is also provided. Substantially pure Glu_GR of at least about 50% is preferred, at least about 70-80% more preferred, and 95-99% or more homogeneity most preferred, particularly for pharmaceutical uses. Once purified, partially or to homogeneity, as desired, the recombinant Glu_GR or native Glu_GR may then be used to generate antibodies, in assay procedures, etc.

In another aspect, the invention concerns polypeptides and fragments of Gluck. Polypeptides and fragments of Gluck may be isolated from recombinant expression systems or may be synthesized by the solid phase method of Merrifield, Fed. Proc. 21:412 (1962), Merrifield, J. Am. Chem. Soc. 85:2149 (1963), or Barany and Merrifield, in The Peptides, vol. 2, pp. 1-284 (1979) Academic Press, NY, each of which are incorporated herein by reference, or by use of an automated peptide synthesizer. By "polypeptides" is meant a sequence of at least about 3 amino acids, typically 6 or more, up to 100-200 amino acids or more, including entire proteins. For example, the portion(s) of GlugR proteins which bind ligand may be identified by a variety of methods, such as by treating purified receptor with a protease or a chemical agent to fragment it and determine which fragment is able to bind to labeled glutamate in a ligand blot. Polypeptides may then be synthesized and used as antigen, to inhibit ligand-GlugR interaction, etc. should be und rstood that as used herein, reference to GlugR is meant to include the proteins, polypeptides, and fragm nts thereof unless th context indicates oth rwis .

WO 92/10583 PCT/US91/09422

19

In another aspect, the invention provides means for regulating the Glu_GR-ligand interaction, and thus treating, therapeutically and/or prophylactically, a disorder which can be linked directly or indirectly to a Glu_GR or to its ligands, such as glutamate and other endogenous excitatory amino acids. By virtue of having the receptors of the invention, agonists or antagonists may be identified which stimulate or inhibit the interaction of ligand with a Glu_GR. With either agonists or antagonists the metabolism and reactivity of cells which express the receptor are controlled, thereby providing a means to abate or in some instances prevent the disease of interest.

10

15

20.

25

30

35

Thus, the invention provides screening procedures for identifying agonists or antagonists of events mediated by the ligand-Glu_cR interaction. Such screening assays may employ a wide variety of formats, depending to some extent on which aspect of the ligand/receptor/G protein interaction is targeted. For example, such assays may be designed to identify compounds which bind to the receptor and thereby block or inhibit interaction of the receptor with the ligand. Other assays can be designed to identify compounds which can substitute for ligand and therefore stimulate Glu_cR-mediated intracellular pathways. Yet other assays can be used to identify compounds which inhibit or facilitate the association of Glu_cR to G protein and thereby mediate the cellular response to Glu_cR ligand.

In one functional screening assay, the initiation of fertilization activation events are monitored in eggs which have been injected with, e.g., mRNA which codes for GlugR and subsequently exposed to selected compounds which are being screened, in conjunction with or apart from an appropriate ligand. See generally, Kline et al., Science 241:464-467 (1988), incorporated herein by reference. Occytes injected with mRNA coding for GlugR can also b assayed by measurement of free cytosolic Ca²⁺ as describ d above.

10

15

20

25

30

35.

Anoth r scre ning assay is based on the use of mammalian cell lines which express Gluck functionally coupled to a mammalian G protein. In this assay, compounds are screened for their relative affinity as receptor agonists or antagonists by comparing the relative receptor occupancy to the extent of ligand induced stimulation or inhibition of second messenger metabolism. For example, activation of phospholipase C leads to increased inositol monophosphate metabolism. Means for measuring inositol monophosphate metabolism are generally described in Subers and Nathanson, J. Mol. Cell, Cardiol. 20:131-140 (1988), incorporated herein by As noted previously, receptor subtypes that reference. are coupled to the stimulation or inhibition of the second messenger adenylate cyclase can be used in assay systems wherein reporter gene (e.g. luciferase) activity is linked to receptor-ligand interactions.

The screening procedure can be used to identify reagents such as antibodies which specifically bind to the receptors and substantially affect their interaction with ligand, for example. The antibodies may be monoclonal or polyclonal, in the form of antiserum or monospecific antibodies, such as purified antiserum or monoclonal antibodies or mixtures thereof. For administration to humans, e.g., as a component of a composition for in vivo diagnosis or imaging, the antibodies are preferably substantially human to minimize immunogenicity and are in substantially pure form. By substantially human is meant generally containing at least about 70% human antibody sequence, preferably at least about 80% human, and most preferably at least about 90-95% or more of a human antibody sequence to minimize immunogenicity in humans.

Antibodies which bind GlugR may be produced by a variety of means. The production of non-human antisera or monoclonal antibodies, e.g., murin, lagomorpha, equine, etc. is well known and may be accomplished by, for xample, immunizing the animal with the receptor

15

20

25

30

35

mol cule or a preparation containing a desired portion of the receptor molecule, such as that domain or domains which contributes to ligand binding. Receptor subtypespecific antibodies can be generated by immunizing With specific peptides. Small peptides (e.g., about 14-20 amino acids) can be coupled to keyhole limpet hemocyanin, for example, to enhance immunogenicity. For the production of monoclonal antibodies, antibody producing cells obtained from immunized animals are immortalized and screened, or screened first for the production of antibody which binds to the receptor protein and then immortalized. As the generation of human monoclonal antibodies to human Glu R antigen may be difficult with conventional techniques, it may be desirable to transfer antigen binding regions of the non-human antibodies, e.g. the F(ab'), or hypervariable regions, to human constant regions (Fc) or framework regions by recombinant DNA techniques to produce substantially human molecules. Such methods are generally known in the art and are described in, for example, U.S. Patent No. 4,816,397 and EP publications 173,494 and 239,400, which are incorporated herein by reference. Alternatively, one may isolate DNA sequences which code for a human monoclonal antibody or portions thereof that specifically bind to the human receptor protein by screening a DNA library from human B cells according to the general protocol outlined by Huse et al., Science 246:1275-1281 (1989), incorporated herein by reference, and then cloning and amplifying the sequences which encode the antibody (or binding fragment) of the desired specificity.

In other embodiments, the invention provides screening assays conducted in vitro with cells which express the receptor. For example, the DNA which encodes the receptor or selected portions thereof may be transfected into an stablished cell line, .g., a mammalian cell line such as BHK or CHO, using procedures known in the art (see, e.g., Sambrook et al., Molecular Cloning, A Laboratory Manual, 2d ed., Cold Spring Harbor

15

20

25

30

35

Laboratory Pr ss, Cold Spring Harbor, N.Y., 1989, which is incorporated her in by refer nce). The receptor is then expressed by the cultured cells, and selected agents are screened for the desired effect on the cell, separately or in conjunction with an appropriate ligand such as glutamate or quisqualate. Means for amplifying nucleic acid sequences which may be employed to amplify sequences encoding the receptor or portions thereof are described in U.S. Pat. Nos. 4,683,195 and 4,683,202, incorporated herein by reference.

In yet another aspect, the screening assays provided by the invention relate to transgenic mammals whose germ cells and somatic cells contain a nucleotide sequence encoding Glu_GR protein or a selected portion of the receptor which, e.g., binds ligand, GTP binding protein, or the like. There are several means by which a sequence encoding, for example, the human Glu_GR may be introduced into a non-human mammalian embryo, some of which are described in, e.g., U.S. Patent No. 4,736,866, Jaenisch, Science 240-1468-1474 (1988) and Westphal et al., Annu. Rev. Cell Biol. 5:181-196 (1989), which are incorporated herein by reference. The animal's cells then express the receptor and thus may be used as a convenient model for testing or screening selected agonists or antagonists.

In another aspect the invention concerns diagnostic methods and compositions. By means of having the Gluck molecule and antibodies thereto, a variety of diagnostic assays are provided. For example, with antibodies, including monoclonal antibodies, to Gluck, the presence and/or concentration of receptor in selected cells or tissues in an individual or culture of interest may be determined. These assays can be used in the diagnosis and/or treatment of diseases such as, for xample, cerebral ischemia, Parkinsons, senile dementia and other cognitive disorders, Huntington's chorea, amyotrophic lateral sclerosis, emesis, migraine, and others.

10

15

20

25

30

35

Numerous types of immunoassays are availabl and are known to those skilled in the art, .g., competitive assays, sandwich assays, and the like, as generally described in, e.g., U.S. Pat. Nos. 4,642,285; 4,376,110; 4,016,043; 3,879,262; 3,852,157; 3,850,752; 3,839,153; 3,791,932; and Harlow and Lane, Antibodies, A Laboratory Manual, Cold Spring Harbor Publications, N.Y. (1988), each incorporated by reference herein. In one assay format Gluck is identified and/or quantified by using labeled antibodies, preferably monoclonal antibodies which are reacted with brain tissues, e.g., cortex, striatum, hippocampus, cerebellum, and determining the specific binding thereto, the assay typically being performed under conditions conducive to immune complex formation. Unlabeled primary antibody can be used in combination with labels that are reactive with primary antibody to detect the receptor. For example, the primary antibody may be detected indirectly by a labeled secondary antibody made to specifically detect the primary antibody. Alternatively, the anti-GlugR antibody can be directly labeled. A wide variety of labels may be employed, such as radionuclides, particles (e.g., gold, ferritin, magnetic particles, red blood cells), fluorophores, chemiluminescers, enzymes, enzyme substrates, enzyme cofactors, enzyme inhibitors, ligands (particularly haptens), etc.

The Glu_CR DNA may be directly detected in cells with a labeled Glu_CR DNA or synthetic oligonucleotide probe in a hybridization procedure similar to the Southern or dot blot. Also, the polymerase chain reaction (Saiki et al., <u>Science</u> 239:487 (1988), and U.S. Pat. No. 4,683,195) may be used to amplify DNA sequences, which are subsequently detected by their characteristic size on agarose gels, Southern blots of these gels using Glu_CR DNA or a cligonucl otide prob, or a dot blot using similar probes. The probes may c mprise from about 14 nucleotides to about 25 or more nucleotides, preferably, 40 to 60 nucleotid s, and in some instanc s a substantial

.10

15

20

25

30

portion or even the ntire cDNA f Glu_GR may be used. The prob s ar labeled with a detectable signal, such as an enzyme, biotin, a radionuclide, fluorophore, chemiluminescer, paramagnetic particle, etc.

Kits can also be supplied for use with the receptor of the subject invention in the detection of the presence of the receptor or antibodies thereto, as might be desired in the case of autoimmune disease. Thus, antibodies to Gluck, preferably monospecific antibodies such as monoclonal antibodies, or compositions of the receptor may be provided, usually in lyophilized form in a container, either segregated or in conjunction with additional reagents, such as anti-antibodies, labels, gene probes, polymerase chain reaction primers and

polymerase, and the like.

The following examples are offered by way of illustration, not by limitation.

EXAMPLE I

Preparation of Gluck enriched mRNA

Total RNA was prepared from the cerebellum of rats using guanidine isothiocyanate (Chirgwin et al. Biochemistry 18:52-94 (1979)) and CsCl centrifugation. Poly(A)+ RNA was isolated using oligo d(T) cellulose chromatography. After 2 rounds of chromatography on oligo d(T) cellulose the RNA (800 μ g) was divided into two aliquots and layered over 10-40% linear sucrose gradients in tubes for an SW 28 rotor. The gradients were centrifuged for 28 hours at 25,000 rpm to pellet RNA greater than 4 kb in size. The enriched RNA was injected into frog occytes and assayed for the presence of the Glu_GR.

15

20

25

30

35

Injecti n of occytes and voltage-clamp assay of GlugR activity

Occytes were prepared from ovarian lobes that were surgically removed from anesthetized Xenopus females. The ovarian lobes were washed, pulled apart into small clumps and dissociated by treatment with collagenase for 2-3 hours at 20°C with constant, gentle agitation. The dissociation and defolicularization of the occytes is completed manually after removal of the collagenase. Occytes that were judged healthy and greater than 1 mm in diameter were transferred to a 50 mm sterile tissue culture dish and incubated in sterile, antibiotic-supplemented Barth's medium (88 mM NaCl, 1mM KCl, 0.82 mM MgSO₄, 0.33 mM Ca(NO₃)₂, 0.41 mM CaCl₂, 2.4 mM NaHCO₃, 10 mM HEPES, pH 7.4, 0.1 mg/ml gentamicin, 0.01 mg/ml penicillin, 0.01 mg/ml streptomycin, 0.5 mM theophylline, and 2.5 mM Na pyruvate) at 19°C.

Injection pipettes were pulled from hard glass tubing (Drummond) on a modified 700C Kopf vertical puller. The tip was broken and bevelled using a List Medical microforge. Tip diameters of the pipettes ranged from 20-30 mM. Injection pipettes were made RNase free by heating to 285°C overnight.

Following overnight incubation, healthy oocytes were selected for injection. RNA, which was stored at -70°C in DEPC-treated water, was thawed and centrifuged at 15,000 g for five minutes. Injection was performed using a modified pipetting device (Drummond). After injection, the oocytes were incubated in fresh, sterile Barth's medium which was changed daily, and unhealthy oocytes were removed.

Voltage-clamp assays were carried out on injected oocytes which were each placed in a small chamber of approximately 500 µl in volume and which was continuously p rfused with standard frog Ringer's (115 mM NaCl, 2.5 mM KCl, 1.8 mM CaCl, 10 mM HEPES, pH 7.2) at 1-6 ml/min. The oocyte was impaled with two glass microelectrodes for recording which, when filled with 3 M

KCl, had a tip resistanc of 0.5 to 7.0 megaohms. One of the two lectrodes was connect d to a diff rential amplifier via a silver/silver chloride half cell. The bath potential was measured by connecting the other side of the differential amplifier to the bath via a silver/silver chloride pellet and a Ringer/Agar bridge. A low noise, high compliance, voltage-clamp system (NPI) was used to control the membrane potential and to measure membrane current. The oocyte membrane potential was maintained at -60 mV (inside cell negative). millimolar glutamate (Sigma), 100 μ M quisqualate (Sigma), 1 mM carbamylcholine (Sigma), and the other drugs used in this assay were applied by switching the perfusing medium to a medium containing a drug for approximately three minutes, and the membrane current was recorded on a chart recorder (Linear Instruments).

After impaling the oocyte with the two microelectrodes, and imposing the voltage-clamp, the membrane current (the holding current) gradually declines to a steady state over a period of several minutes. When the holding current stabilizes, so that the chart record is horizontal, the drug is applied for one to three minutes. An oocyte is judged to have a positive response if a rapid inward current spike (downward deflection on the chart), followed by slow current oscillations of decreasing magnitude, is observed. Our lower limit of detection depended on the steadiness of the holding current prior to drug application, but was in the range of 5-10 nA.

30

35

10

15

20

25

Construction of pVEGT'

To permit transcription of cloned cDNA without prior endonuclease digestion, bacteriophage T7 transcriptional terminators were added to a cloning vector. Plasmid pVEGT' is described in copending U.S.S.N. 07/581,342, which is incorporated by reference her in. The sequence of the putative T7 RNA transcription t rminator, which lies between gene 10 and

15

20

25

30

35

gene 11 of bacteri phag T7, is disclosed by Dunn and Studier (J. Mol. Biol. 166: 477-536 (1983)). As shown in Figure 5, four synthetic oligonucleotides were designed from this sequence and ligated into the vector pGEM-1 (obtained from Promega Biotec, Madison, WI), a plasmid containing a bacterial origin of replication, ampicillin resistance gene, and the T7 promoter adjacent to a multiple cloning site. Terminal phosphates were added to the 5' ends of oligonucleotides ZC776 and ZC777 (Sequence ID Nos. 4 and 5) with T4 polynucleotide kinase and ATP, under standard conditions (Maniatis et al. ibid). (The sequences of these and other oligonucleotides referred to herein are shown in Table 1.) After the incubation, the kinase was heat killed at 65°C for 10 min. Twenty-five nanograms of oligonucleotide ZC775 (Sequence ID Number 3) and 25 ng of oligonucleotide ZC776 (Sequence ID Number 4) were annealed by incubation at 65°C for 15 minutes, then allowed to cool to room temperature in 500 ml of water. Oligonucleotides ZC777 and ZC778 (Sequence ID Nos. 5 and The annealed 6) were similarly annealed. oligonucleotides were stored at -20°C until use. The vector pGEM-1 was digested with Pst I and Hind III, and the linearized vector DNA was purified by agarose gel electrophoresis. The synthetic T7 terminator (annealed oligonucleotides ZC775, ZC776, ZC777 and ZC778; Sequence ID Nos. 3, 4, 5 and 6) was then cloned into pGEM-1. Twenty-five nanograms of vector plus an equal molar amount of each of the annealed oligonucleotides ZC775/ZC776 (Sequence ID Nos. 3 and 4) and ZC777/ZC778 (Sequence ID Nos. 5 and 6) were combined in a 10 μ l reaction mix. After an overnight ligation at 14°C, the DNA was transformed into competent E. coli JM83 cells, and the transformed cells were selected for ampicillin resistance. Plasmid DNA was prepared from selected transformants by the alkaline lysis procedure (Birnboim and Doly, Nuc. Acids Res. 7:1513-1523 (1979)). A portion of th DNA from these samples was cut with Pst I and Hind III and analyz d on a 4% polyacrylamid gel to id ntify

15

25

30

35

cl nes that released an 80 bp Pst I-Hind III fragment. Other diagnostic cuts, such as Eco RI and Not I, were also made. One of the isolates, designated pGEMT, was shown by restriction analysis to contain the T7 terminator fragment.

Table 1

Oligonucleotide Sequences (5' - 3')

ZC775 (Sequence ID Number 3):
GCT AGC ATA ACC CCT TGG GGC CTC TAA ACG GGT CT

ZC776 (Sequence ID Number 4): CTC AAG ACC CGT TTA GAG GCC CCA AGG GGT TAT GCT AGC TGC A

ZC777 (Sequence ID Number 5):

TGA GGG GTT TTT TGC TGA AAG GAG GAA CTA TGC GGC CGC A

ZC778 (Sequence ID Number 6):

AGC TTG CGG CCG CAT AGT TCC TCC TTT CAG CAA AAA ACC C

ZC1751 (Sequence ID Number 7):

AAT TCT GTG CTC TGT CAA G

ZC1752 (Sequence ID Number 8):

GAT CCT TGA CAG AGC ACA G

ZC2063 (Sequence ID Number 9):

GAT CCA AAC TAG TAA AAG AGC T

ZC2064 (Sequenc ID Number 10):

CTT TTA CTA GTT TG

(Table 1, continued)

ZC2938 (Sequence ID Number 11):

GAC AGA GCA CAG ATT CAC TAG TGA GCT CTT TTT TTT TTT T

ZC3015 (Sequence ID Number 12):

TTC CAT GGC ACC GTC AAG GCT

10

5

ZC3016 (Sequence ID Number 13):

15 AGT GAT GGC ATG GAC TGT GGT

ZC3652 (Sequence ID Number 14):

ACA TGC ACC ATG CTC TGT GT

ZC3654 (Sequence ID Number 15):

AGT GAT GGC ATG GAC TGT GGT

25

30

35

40

20

The native T7 terminator from plasmid pAR2529
(Rosenberg et al., Gene 56:125-135 (1987)) was added to plasmid pGEMT. Plasmid pGEMT was digested with Bam HI and plasmid pAR2529 was digested with Bam HI and Bgl II
(Figure 1). The Bam HI-Bgl II terminator fragment from pAR2529 was purified by agarose gel electrophoresis. The terminator fragment was ligated to Bam HI digested pGEMT, and the DNA was transformed into competent E. coli LM1035 cells. Colonies that were ampicillin resistant were inoculated into 5 ml cultures for overnight growth. Plasmid DNA prepared by the alkaline lysis procedure was screened for proper terminator orientation by Bam HI-Sal I dig stion and electrophoresis on an 8% polyacrylamide g 1. A clon that contain d the terminator in th correct ori ntation, as evidenced by the pres nce of a

15 .

20

25

30

35

130 bp Bam HI-Sal I fragment, was chos n and named pGEMTT (Figure 1).

To allow pGEMTT to be packaged as single-stranded DNA in the presence of M13 phage proteins, the M13 intergenic region from pUC382 (similar to pUC118 and 119 as disclosed by Vieira and Messing, Methods Enzymol. 153: 3-11 (1987) was added to pGEMTT (Figure 1). Plasmid pGEMTT was digested with Fsp I and Nar I, and the fragment containing the T7 promoter and transcription terminator was purified. Plasmid pUC382 was digested with Fsp I and Nar I, and the fragment encoding the ampicillin resistance gene and the M13 intergenic region was gel purified. These fragments were then ligated together in the presence of T4 DNA ligase. DNA was transformed into competent E. coli LM1035 cells. Plasmid DNA from twelve ampicillin-resistant colonies was prepared by the alkaline lysis method, and the DNA was screened by digestion with Ava I. The appropriate construction gave two bands, one of 2430 bp and another of 709 bp. One such isolate was chosen and named pVEG. Synthetic oligonucleotides encoding the prime sequence were added to pVEG between the Bam HI and Eco RI sites (Figure 1). Plasmid pVEG was digested with Bam HI and Eco RI and the vector fragment was gel purified. Ninety-six nanograms each of oligonucleotides ZC1751 and ZC1752 (Sequence ID Nos. 7 and 8) were annealed in 4.5 μ l of 10 mM Tris pH 7.5, 20 mM MgCl, and 10 mM NaCl at 65°C for 20 minutes, then the mixture was cooled to room temperature over a period of 30 minutes. The annealed oligonucleotides were ligated to the pVEG vector fragment with T4 DNA ligase and then transformed into competent E. coli LM1035 cells. After growing overnight to develop the colonies, a filter lift was taken of the colonies on the agar plate. The filter was probed with 32P-labeled oligonucleotid ZC1751 (Sequ nc ID Number 7). th colonies w re positive. Plasmid DNA was pr pared from cultures grown from 12 of the colonies. The plasmid DNA was scr ened by digestion with Sst I to verify the

10

15

20

25

30

35

absence of the Sst I site between the Eco RI and Bam HI sites of pVEG. All 12 of the plasmid DNAs were negative for Sst I digestion. One of these 12 isolates was chosen and named pVEG'.

A polyadenylate sequence derived from an Aspergillus alcohol dehydrogenase cDNA was added to pVEG. As shown in Figure 1, plasmid pM098 (disclosed in published European patent application EP 272,277 and deposited with American Type Culture Collection under accession number 53428) was digested with Dra I and Bam HI, and the approximately 150 bp poly(A) fragment was purified by agarose gel electrophoresis. This fragment contained mostly poly(A) sequence with very little flanking cDNA. To clone the poly(A) cDNA fragment into pVEG, pVEG was digested with Bam HI and Sma I, and the 3.4 kb vector fragment was gel purified. The vector and poly(A) fragments were ligated together with T4 DNA ligase to produce vector pVEGT (Figure 1).

Synthetic oligonucleotides encoding the prime sequence were added to pVEGT. To accomplish this, pVEGT was digested with Not I and Sst I, and the 370 bp fragment containing the poly(A) sequence and the two T7 transcriptional terminators was purified by agarose gel electrophoresis. Plasmid pVEG' was digested with Not I and Bam HI, and the 3.2 kb vector fragment was gel-purified. Two oligonucleotides (ZC2063 and ZC2064; Sequence ID Nos. 9 and 10) that formed, when annealed, a Bam HI-Sst I adapter were synthesized. The two oligonucleotides were individually kinased and annealed, and ligated with the linearized vector and the poly(A)-terminator fragment. The resultant vector, designated pVEGT' (Figure 1), contained a T7 RNA transcription promoter, an Eco RI cloning site flanked by the prime sequence, a poly(A) tract, and two T7 RNA polymerase terminators.

15

20

25

30

35

Construction of cDNA library from rat cerebellum poly(A)+
RNA

Because there was evidence suggesting that the Glu R was encoded a very large mRNA of 7 kb (Fong, Davidson, and Lester, Synapse 2:657 (1988)) and because full length cDNA encompassing the coding sequence is required for functional cloning of cDNA, measures were taken to optimize for synthesis of large cDNA. A novel method of cDNA synthesis was developed which yielded large full length cDNA. This was evident by demonstration that full length 7.5 kb cDNA could be synthesized from a model 7.5 kb mRNA and that large full length cDNA were present in a library constructed from poly(A) + RNA as demonstrated by Southern blot analysis. In addition, all enzymes which were important in this method were pretested and selected from a large number of lots of enzymes available from commercial suppliers. Once a satisfactory lot was identified, a large amount of the enzyme was purchased and the enzyme was stored at -70°C until used. Once used, the enzyme was stored at -20°C for a few months and then discarded. Different "lots" of enzymes from commercial suppliers, including lots of Superscript reverse transcriptase (BRL), E. coli DNA polymerase I (Amersham) and Mung bean nuclease (NEB), which were used in the cDNA synthesis, were screened for quality in test synthesis assays. Superscript reverse transcriptase lots were assayed for the ability to synthesize unit length (7.5 kb) first strand cDNA from 7.5 kb RNA (BRL) control. Conditions for first strand synthesis with Superscript reverse transcriptase lots were prepared as described below. Radiolabeled first strand cDNA was analyzed by alkaline agarose gel electrophoresis. Superscript lots capable of producing unit length, 7.5 kb cDNA were selected for use.

E. coli DNA polymeras I lots w re assayed for the ability t produce, by hairpin DNA formation, full-length second strand cDNA from the 7.5 kb unit-length first strand cDNA. The second strand cDNA syntheses were

10

15

20

25

30

35

carried out as described below. The quality f the second strand syntheses were assessed by alkaline agarose electrophoresis of the radiolabeled product. DNA polymerase I lots capable of producing 15 kb second strand DNA from the 7.5 kb unit length first strand cDNA were selected for use.

Mung bean nuclease lots were tested for the ability to clip the hairpin DNA formed during second strand synthesis without degrading the cDNA. In addition, varying concentrations of enzyme were added to determine the optimum enzyme concentration for the conditions set forth below. The reactions were assessed by alkaline agarose electrophoresis. Lots and concentrations resulting in the production of 7.5 kb unit length cDNA were selected for use.

Total RNA was prepared from rat cerebella using guanidine isothiocyanate (Chirgwin et al. <u>Biochemistry</u> 18:52-94 1979) and CsCl centrifugation (Gilsin et al. <u>Biochemistry</u> 13:2633-2637 1974). Poly(A)+ RNA was selected from the total RNA using oligo d(T) cellulose chromatography (Aviv and Leder, <u>Proc. Natl. Acad. Sci. USA</u> 69:1408 (1972)).

First strand cDNA was synthesized from one time poly d(T)-selected cerebellum poly(A)+ RNA in two separate reactions. One reaction, containing radiolabeled dATP, was used to assess the quality of first strand synthesis. The second reaction was carried out in the absence of radiolabeled dATP and was used, in part, to assess the quality of second strand synthesis. Superscript reverse transcriptase (BRL) was used specifically as described below. A 2.5x reaction mix was prepared at room temperature by mixing, in order, 10 μ l of 5x reverse transcriptase buffer (BRL; 250 mM Tris-HCl pH 8.3, 375 mM KCl, and 15 mM MgCl₂), 2.5 μ l 200 mM dithiothreitol (made fresh or stored in aliquots at -70°C) and 2.5 μ l of a d oxynucleotide triphosphate solution containing 10 mM each f dATP, dGTP, dTTP and 5-methyl dCTP (Pharmacia). The reaction mix was

10

15

20

25

30

35

aliquoted into two tub s of 7.5 μ l ach. To the first tube, 1.3 μ l of 10 μ Ci/ μ l α^{32} P-dATP (Am rsham) was added and 1.3 μ l of water was added to the second reaction Seven microliters from each tube was transferred to reaction tubes. Fourteen microliters of a solution containing 10 μ g of cerebellum poly(A)+ RNA diluted in 14 μ l of 5 mM Tris-HCl pH 7.4, 50 μ M EDTA was mixed with 2 μ l of 1 μ g/ μ l first strand primer, ZC2938 (Table 1; Sequence ID No. 11), and the primer was annealed to the RNA by heating the mixture to 65°C for 4 minutes, followed by chilling in ice water. Eight microliters of the RNA-primer mixture was added to each of the two reaction tubes followed by 5 μ l of 200 U/ μ l Superscript reverse transcriptase (BRL). The reactions were mixed gently, and the tubes were incubated at 45°C for 30 minutes. After incubation, 80 μ l of 10 mM Tris-HCl pH 7.4, 1 mM EDTA was added to each tube, the samples were vortexed and centrifuged briefly. Three microliters of each reaction was removed to determine total counts and TCA precipitable counts (incorporated counts). Two microliters of each sample was analyzed by alkaline gel electrophoresis to assess the quality of first strand synthesis. The remainder of each sample was ethanol precipitated. The nucleic acids were pelleted by centrifugation, washed with 80% ethanol and air dried for ten minutes. The first strand synthesis yielded 1.4 μg of cerebellum cDNA or a 28% conversion of RNA into DNA.

Second strand cDNA synthesis was performed on the RNA-DNA hybrid from the first strand reactions under conditions which encouraged first strand priming of second strand synthesis resulting in DNA hairpin formation. The nucleic acid pellets containing the first strand cDNA were resuspended in 71 μ l of water. To assess the quality of second strand synthesis, $\alpha^{32}P$ -dATP was added to the unlabeled first strand cDNA. To ncourage formation of the hairpin structur, all reag nts except the enzymes were brought to room temperature, and the reaction mixtur s were set up at

10

15

20

25

30

35

room temperature. (Alternatively, the reagents can be on ice and th reaction mixture set up at room t mp rature and allowed to equilibrate at room temperature for a short time prior to incubation at 16°C.) Two reaction tubes were set up for each synthesis. One reaction tube contained the unlabeled first strand cDNA and the other reaction tube contained the radiolabeled first strand To each reaction tube, 20 μ l of 5x second strand CDNA. buffer (100 mM Tris, pH 7.4, 450 mM KCl, 23 mM MgCl₂, 50 mM (NH₄)₂SO₄), 3 μ l of beta-NAD and 1 μ l of a deoxynucleotide triphosphate solution containing 10 mM each of dATP, dGTP, dTTP and dCTP (Pharmacia), 1 μ l α^{32} P-dATP or 1 μ l of water (the radiolabeled dATP was added to the tube containing the unlabeled first strand cDNA), 0.6 μ l of 7 U/ μ l E. coli DNA ligase (Boehringer-Mannheim), 3.1 μ l of 8 U/ μ l E. coli DNA polymerase I (Amersham), and 1 μ l of 2 U/ μ l of RNase H The reactions were incubated at 16°C for 2 hours. After incubation, 3 μ l was taken from each reaction tube to determine total and TCA precipitable counts. microliters of each sample was analyzed by alkaline gel electrophoresis to assess the quality of second strand synthesis by the presence of a band of approximately twice unit length. To the remainder of each sample, 2 μ l of 2.5 μ g/ μ l oyster glycogen, 5 μ l of 0.5 M EDTA and 200 μ l of 10 mM Tris-HCl pH 7.4, 1 mM EDTA were added, the samples were phenol-chloroform extracted, and isopropanol precipitated. The nucleic acids were pelleted by centrifugation, washed with 80% ethanol and air dried. The yield of double stranded cDNA in each of the reactions was approximately 2 μ g.

The single-stranded DNA in the hairpin structure was clipped using mung bean nuclease. Each second strand DNA sample was resuspended in 12 μ l of water. Two microliters of 10x mung bean buffer (0.3 M NaOAC, pH 4.6, 3 M NaCl, 10 mM ZnSO₄), 2 μ l of 10 mM dithi threit 1, 2 μ l of 50% glycer 1, and 2 μ l of 10 U/ μ l mung b an nuclease (NEB, lot 7) wer add d to each tube, and the reactions

w r incubated at 30°C for 30 minutes. After incubation, 80 μ l of 10 mM Tris-HCl pH 7.4, 1 mM EDTA was added to each tube, and 2 μ l of each sample was subjected to alkaline gel electrophoresis to assess the cleavage of the second strand product into unit length cDNA. One hundred microliters of 1 M Tris-HCl pH 7.4 was added to each sample, and the samples were twice extracted with phenol-chloroform. Following the final phenol-chloroform extraction, the DNA was isopropanol precipitated. The DNA was pelleted by centrifugation, washed with 80% ethanol and air dried. Approximately 2 μ g of DNA was obtained from each reaction.

The cDNA was blunt-ended with T4 DNA polymerase after the cDNA pellets were resuspended in 12 μ l of water. Two microliters of 10x T4 DNA polymerase buffer (330 mM Tris-acetate, pH 7.9, 670 mM KAc, 100 mM MgAc, 1 mg/ml gelatin), 2 μ l of 1 mM dNTP, 2 μ l 50 mM dithiothreitol, and 2 μ l of 1 U/ μ l T4 DNA polymerase (Boehringer-Mannheim) were added to each tube. After an incubation at 15°C for 1 hour, 180 μ l of 10 mM Tris-HCl pH 7.4, 1 mM EDTA was added to each sample, and the samples were phenol-chloroform extracted followed by isopropanol precipitation. The cDNA was pelleted by centrifugation, washed with 80% ethanol and air dried. Eco RI adapters (Invitrogen, Cat. # N409-20) were ligated to the blunted cDNA after the DNA from each reaction was resuspended in 6.5 μ l water.

The first strand primer encoded an Sst I cloning site to allow the cDNA to be directionally cloned into an expression vector. The cDNA was digested with Sst I followed by phenol-chloroform extraction and isopropanol precipitation. After digestion, the cDNA was electrophoresed in a 0.8% low melt agarose gel, and the cDNA over 4.2 kb was electroeluted using an Elutrap (Schleicher and Schu 11, Keene, NH). The electroeluted cDNA in 500 μ l of buff r was isopropanol precipitat d and the cDNA was pelleted by centrifugation. The cDNA pell t was washed with 80% thanol.

10

15

20

25

30

35

A cerebellum cDNA library was stablished by ligating th cDNA to th Eco RI-Sst I digested, agarose gel purified pVEGT'.

Ten sublibraries of one million clones each were constructed representing a library of ten million independent clones. To prepare each sublibrary, 80 ng of linearized vector were ligated to 40 ng of cDNA. After incubation at room temperature for 11 hours, 2.5 μ g of oyster glycogen and 80 μ l of 10 mM Tris-HCl, 1 mM EDTA was added and the sample was phenol-chloroform extracted followed by ethanol precipitation. The DNA was pelleted by centrifugation, and the DNA pellet washed with 80% ethanol. After air drying, the DNA was resuspended in 3 Thirty-seven microliters of μ l of water. electroporation-competent DH10B cells (BRL) was added to the DNA and electroporation was completed using a BioRad electroporation unit. After electroporation, 4 ml of SOC (Maniatis et al.) was added to the cells, and 400 μ l was spread on each of 10-150 mm LB ampicillin plates. plate represented a sublibrary of 100,000 clones. After an overnight incubation, the cells were harvested by adding 10 ml of LB ampicillin media to each plate and scraping the cells into the media. Glycerol stocks and plasmid DNA were prepared from each plate. The library background (vector without insert) was established at about 15%.

Detection of Glu R activity from the cDNA library

The <u>Xenopus</u> oocyte efficiently translates exogenously added mRNA. Preliminary experiments were done using the mouse m1 muscarinic receptor cDNA (a G protein-coupled receptor that can be detected by voltage-clamp) cloned into pVEGT'. Injection of RNA transcribed in <u>vitro</u> from increasing dilutions of the m1 template DNA indicated that m1 ag nist induced activity could be det cted f r ne cl n in a pool size of 100,000. A cerebellum sublibrary was plated into ten p ols of 100,000 unique clones.

15

20

25

30

35

The pools could also be replica plated onto a nitrocellulose filt r and th original and r plica allowed to grow for a few hours. The original plate is scraped to harvest all the colonies. Plasmid DNA is prepared and purified by cesium chloride gradient ultracentrifugation. The DNA from each pool is transcribed in vitro with T7 RNA polymerase in the presence of 7-methyl-G, the capped nucleotide, to increase translation efficiency. Template DNA transcription reactions are spiked with a dilution of two control genes cloned into pVEGT': the mouse ml gene and a secreted version of the human placental alkaline phosphatase gene (SEAP; Tate et al., Fed. Am. Soc. Exp. Biol. 8: 227-231 (1990), incorporated by reference herein). Transcription from the control genes would allow selection of occytes that more efficiently translate the injected RNA, and a determination whether oocytes that are negative for the GlugR are true negatives, that is, still having a detectable ml agonistinduced response.

Plasmid DNA prepared from each of the 10 pools of 100,000 clones, which in total represented one sublibrary of one million clones of the cerebellum cDNA library, was purified by cesium chloride gradient ultracentrifugation. The DNA was transcribed in vitro with T7 RNA polymerase (Pharmacia) in the presence of capped nucleotide (GpppG, Pharamcia). The presence of a poly(A) sequence and two T7 RNA polymerase terminators in pVEGT' resulted in RNA with a capped 5' end, the sequence of the cDNA insert, and 3' poly(A) tails. Capped RNA is believed necessary for efficient translation in oocytes (Noma et al. Nature 319:640 (1986)) and the poly(A) sequence has been shown to increase the synthesis of a protein in oocytes by more then 40 fold. The transcription reaction tubes were set up by adding 12 μ l of 5x transcription buffer (Stratag ne Cloning Systems, La J lla, CA), 3 μ l ach of 10 mM ATP, CTP, GTP, and UTP, 6 μ l of 10 mM GpppG (Pharmacia), 6 μ l of 1 mg/ml BSA, 3 μ l of 200 mM DTT, 1.5 μ l of 40 U/ μ l

10

15

20

25 .

30

35

RNasin (ProMega Biotech, Madis n, WI), 8.5 μ l f wat r, 10 μ l of cDNA containing 5 t 10 μ g DNA, and 1 μ l of 70 $U/\mu l$ T7 RNA polymerase. After mixing, 10 μl of the reaction was transferred to a tube containing 0.5 \(\mu Ci \) of α^{32} P-UTP to determine the total counts and counts incorporated into RNA. The samples were incubated at 37°C for one hour. The cDNA in the unlabeled samples was degraded with the addition of 1 μ l of 200 mM DTT, 2 μ l of 30 U/ μ l DNase I, and 0.5 μ l of 40 U/ μ l RNasin and the incubation was continued at 37°C for 15 minutes. Forty microliters of water was added to the radiolabeled reactions, and 1 μ l was removed from each sample and counted to determine total counts. The remainder of the labeled samples were ethanol precipitated. The samples were centrifuged to collect the RNA and the RNA pellets were counted to determine the counts incorporated into After the DNA degradation reaction in the unlabeled RNA. samples, 70 μ l of 10 mM Tris-HCl, 1 mM EDTA was added to each sample, and the samples were twice-extracted with phenol-chloroform followed by one chloroform extraction. The RNA was ethanol precipitated. After centrifugation to collect the RNA, the pellets were washed with 80% ethanol, followed by air drying for 10 minutes. A typical yield of the unlabeled RNA was 20 to 30 μ g. The unlabeled RNA was resuspended at 2 μ g/ μ l in diethylpyrocarbonate (DEPC, Sigma) treated water and stored at -70°C.

Prior to microinjection into oocytes, the RNA samples were thawed and centrifuged in a microfuge for 5 minutes to remove any particles that might clog a microinjection pipet. After centrifugation, 80% of each sample was removed and split into two tubes.

The RNA from each of the 10 sublibraries were injected into oocytes as described above and translation was allow d for four days. Expression of Glu_GR activity was ass ssed by voltag -clamp assay as describ d above. One of the 10 sublibrari s, Z93-1.9, produced a signal with administration of quisqualat to th oocyte.

<u>subdivision</u> of the cDNA library pool to obtain pure Gluck clone

The DNA pool (Z93-1.9) was subdivided by plating clones from the glycerol stock onto LB ampicillin plates. To determine the number of clones that should be plated for the subdivision of the 100,000 clone pool to identify a positive clone, the probability equation N = ln (1- P) / ln (1 - f) (Maniatis et al., ibid.) was used, where P is the desired probability of including the clone of interest, f is the fraction of positive clones in the pool, and N is the number of clones to be plated to provide the given probability. For a probability of 99.8% for a pool size of 100,000 to contain one positive clone, 621,461 clones should be plated.

Forty-eight 150 mm LB ampicillin plates were plated with the glycerol stock representing the 100,000 positive pool, Z93-1.9, at a density of approximately 14,000 clones per plate to give a total of 670,000 clones. After an overnight incubation 37°C, the bacteria on each plate were harvested into 10 ml of Solution I (as described by Birnboim and Doly, Nuc. Acids Res. 7:1513 (1979)), incorporated by reference herein). A glycerol stock was prepared from a portion of the cells, and plasmid DNA was prepared from the remainder of the cells. Six pools of DNA representing eight of the LB ampicillin plates each were prepared by combining one tenth of the plasmid DNA from groups of eight plates into each pool. The plasmid DNA from these six pools was purified by cesium chloride gradient centrifugation. The DNA was transcribed into RNA as outlined above. Transcription of the parent pool Z95-1.9 was included as the positive control. Oocytes were injected with the RNA and voltage-clamp assays on the oocytes identified pool Z99-25-32 as positive for GlugR. Pool Z99-25-32 contained DNA prepar d from plat s 25 through 32.

Plasmid DNA from plates 25 to 32 were cesium chlorid banded and transcribed int RNA as d scribed above al ng with the positiv par nt pool 299-25-32.

15

10

20

25

30

WO 92/10583

10

15

20

25

30

35

Occytes were injected with the RNA and voltage clamp assays, carried out as described abov, identified pools Z104-25 and Z111-32 as being weakly positive, Z106-27 and Z109-30 as intermediately positive, and Z108-29 and Z110-31 as the most positive. The pool resulting in Z110-31 was chosen for further subdivision.

Identification of positive pools from the subdivision of the positive pool of 14,000 (Z110-31) from the glycerol stock was unsuccessful. Therefore, plasmid DNA prepared from the pool resulting in Z110-31 was electroporated into bacteria and plated on 60 plates at a density of 1,000 clones/plate. Plasmid DNA was prepared from the bacteria harvested from each plate. Aliquots of the plasmid DNA from each plate were mixed to make six pools representing ten plates each. The plasmid DNA was cesium chloride banded, and the RNA was transcribed as described above. RNA was transcribed from pools Z108-29, Z110-31, and a muscarinic receptor cDNA, m1, for use as positive controls. The RNA was injected into occytes and voltage-clamp assays were carried out as described above. The assays identified pool Z133-21 to 30 as positive.

Plasmid DNA from plates 21 to 30 were cesium chloride banded and transcribed as described above. The transcribed RNA and the RNA from the parent pool Z133-21 to 30 were injected into oocytes and assayed as described above. The voltage-clamp assay identified pool Z142-22 as positive.

Identification of positive pools by the subdivision of the positive pool Z142-22 from a glycerol stock proved unsuccessful. Restriction analysis of plasmid DNA prepared from randomly selected clones from pools Z110-31 (the pool of 14,000) and Z142-22 (the pool of 1,000) indicated that 50% of pool Z110 - 31 and 68% of pool Z142 - 22 were clones without inserts.

T assess physical methods for enriching for the $Glu_{c}R$ cl ne and to establish how many clones fr m pool Z142-22 need d t be assay d t include a $Glu_{c}R$ clone, undigest d plasmid DNA from po 1 Z142-22 was

10:

15

20

25

30

35

electrophoresed on an agarose gel. The sup r-coil band representing v ctor without insert was cut out and the remainder of the DNA was eluted from the gel. The DNA was then electroporated into bacteria cells, and plated at densities of 3,400, 6,900, and 13,800 clones per plate. The plates were replica plated and grown overnight. Plasmid DNA was prepared from the cells harvested from the replica of each plate. The plasmid DNA was transcribed, and the RNA was assayed in cocytes as described above. As a control, each pool contained the equivalent of one colony of ml as an internal positive control. In addition, ml was used as an external positive control. The voltage-clamp assays identified the DNA from the 6,900 clone pool (Z167-7) as positive.

The clones represented on the 6,900 clone plate that resulted in the positive pool Z167-7 were subdivided by replica plating the master plate onto a Biodyne-A nylon membrane on an LB ampicillin plate. The replica plate was incubated four hours at 37°C. After incubation, sub-pools were prepared by removing the membrane from the plate, taping the membrane to a sterile glass plate on a light box, and overlaying the membrane with a grid which divided the membrane into 100 sections. The sections of the grid and underlying membrane were then cut out with a razor blade that had been dipped in alcohol and flamed between each cut. Alcohol-treated, flamed forceps were used to transfer each membrane section to a test tube containing 12.5 ml of LB ampicillin media. The cultures containing the membrane sections were incubated overnight at 37°C. After incubation, 0.5 ml of each culture was mixed with 0.5 ml of 50% glycerol and stored at -70°C to establish glycerol stocks of each sub-pool. Aliquots of the 100 cultures wer po led in a 10 X 10 matrix with sampl s (1) through (10) on the abscissa and samples (a) through (j) on the ordinate. For exampl , 1 ml of cultur s (1) through (10) were added t tube 1 and 1 ml of cultur s (1), (11),

(21), (31), (41), (51), (61), (71), (81), and (91) were added to tub (a) and so on until 10 rows of 10 and 10 columns containing pools of 10 cultures each were completed. Ten microliters of an overnight culture containing ml-transformed bacteria was added to each pool as an internal control. Plasmid DNA was prepared from the 20 sub-pools, and the DNA was purified by cesium chloride gradient centrifugation. RNA was transcribed from the plasmid DNA and was assayed in occytes as described above. Positive controls were the parent pool Z167-7 and pure ml RNA. The voltage-clamp assays indicated that only pools Z175-1 and Z191-g were positive. Consulting the matrix, this indicated that the membrane section number (7) contained the GlugR clone.

To subdivide the clones contained in section (7), a piece of Biodyne A membrane was applied to the master plate containing section (7), the membrane extending beyond section (7) on each side by half the width of section (7). The membrane was removed from the plate, applied to a fresh LB ampicillin plate colony side up, and incubated overnight at 37°C. The membrane was subdivided as described above with the central region of the membrane, the actual section (7) area, divided into 9 small, equivalent-sized squares and the membrane on each side of section (7) was taken as four additional areas. Each membrane section was used to inoculate a 10 ml liquid culture. Bacteria transformed with the ml clone were used as an internal control in each culture as described above. After overnight incubation at 37°C, plasmid DNA was prepared, and the DNA was purified by cesium chloride gradient centrifugation. RNA was transcribed and assayed in oocytes as described above using RNA from m1 and the parent pool number (7) as positive controls. GlugR activity was found in only pool Z203-7 corresponding to membrane section number (7).

Pool Z203-7 was subdivided by electroporating the plasmid DNA prepared from the membrane section numb r (7) into DH10B electroporation-competent cells. Th

.

5

15

20.

25

30

15

20

25

30

35

transformants were plated at a density enabling individual colonies to be picked. Individuals clones were picked to a master plate and into 2 ml of LB ampicillin media. The cultures were incubated overnight, and plasmid DNA was prepared by the method essentially described by Holms and Quigley (Anal. Bioc. 114: 193, (1981)). Restriction analysis suggested that the clones were grouped into 7 different classes of clones. Plasmid DNA, prepared from each class, representing fifty total clones were prepared, transcribed, and assayed in oocytes as described above. However, none of the clones were positive.

To screen for positive clones, electroporation-competent <u>E</u>. <u>coli</u> DH10B cells were electroporated with the DNA prepared from membrane section number (7) (Z203-7) and were plated at 180, 360, 900, and 1800 colonies per plate. The plates were incubated overnight, and replica plates were prepared as described above. Plasmid DNA prepared from each replica plate was combined with 1 to 1000 parts of ml as an internal control. The DNA pools, the ml clone and the parent pool Z203-7 were transcribed, and the RNA was assayed by oocyte injection. The first transcription and injection showed no positives, however, upon retranscription and reanalysis the 1800 clone pool (Z264-1800) was positive for Glu_GR activity.

To subdivide the positive pool of 1800 (Z264-1800), all of the colonies from the plate of 1800, 1528 in total, were each picked to two 100 mm LB ampicillin agar plates on a 100 colony grid. After overnight growth, one set of the duplicate plates was designated as a master set and was placed at 4°C. The other set was replica plated to a third set of plates. After overnight incubation of these plates, the cells on the replica plates were harvested int media and plasmid DNA was prepared from the pooled cells. As described abov, an internal ml control was included in each DNA preparation. ml DNA and the parent Z264-1800 DNA were

us d as xt rnal positive controls. Plasmid DNA prepar d from the 16 plates was transcribed, and the RNA was assayed in oocytes as described above. One of the pools of 100 clones, Z256-I produced Gluck activity.

To identify which clone of the 100 clones from Z256-I produced the Glu_GR activity, a 10 x 10 matrix of the clones was constructed. A liquid culture of each clone was grown. One milliliter of each culture was added to each of two tubes representing the appropriate row and column of the 10 x 10 matrix. As described previously, plasmid DNA encoding m1 was used as an internal positive control. Plasmid DNA prepared from each tube, m1 DNA and DNA from the parent pool Z264-1800 were transcribed and assayed in oocytes as described above. Glu_GR activity was identified only in row (5) and column (e). Thus, the positive clone number 45 was identified as containing the Glu_GR activity.

10

15

20

25

30

35

To confirm the result, plasmid DNA from clone #45 was prepared, transcribed and assayed in oocytes as described above. The results of the assay indicated that clone #45 was capable of producing Gluck activity. Figure 2 illustrates the data taken from voltage-clamp recordings at several stages in the subfractionation of the cerebellum library. Panel (a) is a recorded response to quisqualate of an oocyte previously injected with in vitro transcribed RNA from a rat cerebellum sublibrary of 100,000 independent colonies; panel (b) shows the response to quisqualate in a cell previously injected with RNA transcribed from a subfractionated pool of 14,000 colonies. The peak current was truncated by the chart recorder, but the actual peak current (estimated from a digital panel meter) was approximately 1300 nA. Panel (c) shows the response to quisqualate in a cell injected with pure GluR RNA from clone 45-A. The amount of RNA injected per occyte was approximately 100 ng, except in pan 1 (c) wher the am unt of RNA was 50 pg.

The following describes an alt rnative means for subdividing and screening a positive pool. Working with

cDNA inserts in a plasmid based rather than a lambdabased vector influences th subfractionation protocol. Once a positive pool is identified, the replica filter is overlayed with another sterile nitrocellulose filter. The filter is cut into 88 pieces by using evenly spaced cuts of 10 rows and 10 columns to form a grid. Each of the 88 pieces is transferred to 10 ml of sterile LB +Amp and grown for several hours. Twenty pools are formed; C 1-10 (corresponding to column number) and R 1-10 (corresponding to row number). An aliquot of each of the 88 subfractions is pipetted into 2 tubes, corresponding to its position in a row and a column. DNA is isolated from the 20 pools, purified on CsCl gradients and transcribed in an in vitro reaction that includes the control m1 and SEAP plasmids. After injection into oocytes and voltage-clamp recording there are 2 positive pools, pinpointing the location of 1 of the 88 original subfractions.

Because the positive clone is still part of a pool it must be further subdivided. The probability equation described above is used to determine the number of clones to be plated for the next subdivision of the pool. The glycerol stock from the positive pool is plated out at, e.g., 3000, 6000 and 18,000 clones per plate. After replica plating the DNA is harvested, transcribed, injected and assayed. The pool which is positive is subdivided into a grid of 88 as described above. The assay is repeated, and a single square of the grid is positive. At the next step of subdivision of the pool, 100 individual colonies to a plate are picked, replica plated, and 20 pools are made for transcription and assay. Positive clones are streaked out, several colonies picked and restriction mapped and template and transcript prepared for injection and assay.

35

30

10

15

20

Characterization of Glu R

To establish that the Gluck encoded by clone 45-A couples to G-protein, clone 45-A Gluck RNA was transcribed and injected into occytes as described above. Two days after injection the oocytes were divided into control and toxin-treated groups. The oocytes in the toxin-treated group were treated with a final concentration of 4 μ g/ml of B. pertussis toxin (List Biological Laboratories Inc., Campbell, CA), and both groups were incubated for 24 hours at 19°C as described by Sugiyama et al., Nature 325:531 (1987) and Moriarity et al., J. Biol. Chem. 264:13521 (1989), both of which are incorporated by reference herein. The oocytes from both the control and toxin-treated groups were subjected to voltage-clamp assays as described previously. In one example, oocytes perfused as described previously with 100 µM L-glutamic acid showed a mean L-glutamic acid-induced current of 264.2 nA +/- 73 nA in control occytes (SEM, n=6) and 57.7 nA +/- 19 nA (n=9) in toxin-treated oocytes. The mean membrane current in the toxin-treated group was significantly smaller (p < 0.01) than in the control group suggesting that oocytes injected with 45-A RNA coupled to a pertussis toxin-sensitive G protein.

L-glutamic acid and some of its structural derivatives that are known to activate Glu_CR currents in a dose-dependent manner were applied to oocytes that had been injected with RNA transcribed from the 45-A clone. RNA was transcribed and oocytes were prepared and injected as previously described. Dose dependent responses were measured using voltage clamp assays were carried out in the presence of increasing concentrations of L-glutamic acid (Sigma), quisqualic acid (Sigma), ibotenic acid (Sigma), or trans 1-amino-cyclopentyl-1,3 dicarboxylic acid (tACPD; Tocris Neuramin, Essex, England). Four or five separate oocyt s w re p rfused with increasing concentrations of a particular drug with 30 minut s b tw n consecutive applications of the drug to minimize any int rf rence from desensitization. Th

25

20

10

30

15

20

25

30

35

responses were normalized to a subsequent response to 100 μ M L-glutamic acid. The data were analyzed using the following equation:

(Fractional current) = (Doseⁿ)/(Doseⁿ) + (EC₅₀)ⁿ, where:

Dose = a dose of drug normalized to that evoked by a subsequent application of 100 μM L-glutamic acid;

Fractional current = the peak current evoked by a dose, as defined above:

 EC_{50} = effective concentration that evokes a 50% response (a measure of the potency of an agonist); and

n = the Hill coefficient, a measure of the cooperativity of the reaction.

Using this equation, the effective concentration at 50% stimulation relative to 100 μ M L-glutamic acid was determined for each dose response experiment. Figure 6 shows a representative dose response curve for varying concentrations of L-glutamic acid. The potency series of glutamate analogs and their associated EC₅₀'s are listed in Table 2.

<u>Table 2</u>

Glutamate Analog Potencies (EC50)

Quisqualic acid 0.681 μ M L-glutamic acid 12.32 μ M Ibotenic acid 32.37 μ M tACPD 376 μ M

In addition, oocytes were exposed to the following L-glutamic acid analogs: aspartic acid (Tocris Neuramin), kainic acid, N-methyl-D-aspartic acid (NMDA; Sigma), 2-amino-4-phosphonobytyric acid (APB; Sigma), c-amino-3-hydroxy-5-methyl-isoxazole-4-propionic acid (AMPA; Research Biochemicals Inc., Wayland, MA) at saturating concentrations and the responses were each normalized to a subsequ nt response to 100- μ M L-glutamate. The L-glutamic acid analogs that were found to be ineffectiv were 1 mM aspartic acid, 1 mM kainic

acid, 100 μ M NMDA + 10 μ M glycine, 100 μ M APB and 100 μ M AMPA.

Voltage clamp assays were also carried out on injected oocytes to measure the inhibition by the putative glutamate G protein-coupled receptor antagonist, 2-amino-3-phosphonopropionic acid (AP3). Voltage clamp assays showed that at 1 mM, DL-AP3 (Sigma) reduced the current evoked by 10 μ M glutamic acid to 59.3 +/- 7.3% of the control.

clone 45 cells were streaked out on LB Amp plates and several colonies were picked, grown up and the DNA isolated. Pure 45-A DNA was prepared and restriction mapped by standard procedures. Clone 45-A has been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD, 20852, under ATCC Accession No. 68497. DNA was digested with single or multiple enzymes. The fragments were separated on both 1% agarose and 4% Nusieve gels by electrophoresis. After electrophoresis the DNA was transferred to nitrocellulose filters using standard protocols for Southern transfer. Restriction sites were mapped based on size and based on

hybridization to Pst I subclones of 45-A DNA. Additionally, the entire 45-A cDNA insert can be isolated by digestion with Not I restriction endonuclease. The Not I insert was kinased with $\gamma^{-32}P$ ATP, and after digestion of half of the sample with Bam HI to remove the 3' label, both samples were subjected to digestion with a number of enzymes known to be present once in the insert. In this way the unique sites could be localized. A restriction map of Glu_GR clone 45-A is shown in Figure 3.

The entire 45-A clone was sequenced in both directions using the dideoxynucleotide chain termination method (Sanger and Coulson, J. Mol. Biol. 94:441 (1975), incorporated herein by reference). Figure 5 (Sequence ID Nos. 1 and 2) shows the DNA sequence and deduced amino acid sequence of clone 45-A. Figure 5 also shows the location of putativ N-link d glycosylation sit s, which

10

15

20

25

30

have been pr dicted to occur at the amino acid sequence Asn-X-Thr.

As shown in Figure 5, seven putative transmembrane domains have been predicted from the deduced amino acid sequence of clone 45-A using the method described by Eisenberg et al. J. Mol. Biol. 179:125-142, (1984), incorporated herein by reference. Only those predicted to be transmembrane multimeric domains were included. additional transmembrane domain (the third) was predicted using the method of Hopp and Woods, Proc. Natl. Acad. Sci. USA 78:3824-3838 (1981). Based on these predictions, the protein encoded by clone 45-A appears to have two unusually large domains on the amino- and carboxy-termini that are not found in any of the other reported G protein-coupled receptors which have the common structural feature of seven predicted membrane spanning regions. Analysis of the deduced amino acid sequence of clone 45-A predicts three other hydrophobic stretches including one at the amino-terminus of the sequence. This amino-terminal hydrophobic stretch may be a signal sequence, although no signal cleavage site is predicted downstream of the sequence.

Poly(A)+ RNA was isolated from total rat brain and rat cerebellum using oligo d(T) cellulose chromatography as described by Aviv and Leder (ibid.). Poly(A)+ RNA from rat retina, rat heart, rat lung, rat liver, rat kidney, rat spleen, rat testis, rat ovary and rat pancreas were purchased from Clonetech. The poly(A)+ RNA samples were analyzed by northern analysis (Thomas, Proc. Natl. Acad. Sci. USA 77:5201-5205 (1980), which is incorporated by reference herein). The RNA was denatured in glyoxal, electrophoresed in agarose and transferred to a nitrocellulose membrane essentially as described by Thomas (ibid.). The northern blot was hybridized with a radiolabeled 3473 bp Eco RI-Xba I fragm nt from the 45-A clone. Autoradiography of the bl t show d hybridization to a major band of approximately 7 kb and a smaller band

15

5

10.

20

25

30

15

20

25

30

35

of approximately 3.8 kb in the total rat brain and rat cerebellum RNA.

Single-stranded cDNA was synthesized using 1 µg of the poly(A) + RNA using Superscript reverse transcriptase (BRL) under conditions described by the manufacturer. One fourth of the cDNA was used as a template for PCR amplification using 40 pmoles each of the GluGR-specific primers ZC3652 (Table 1; Sequence ID Number 14) and ZC3654 (Table 1; Sequence ID Number 15) and 2.5 U Taq I polymerase (Perkin Elmer Cetus, Norwalk, VA) and conditions specified by the manufacturer. As an internal control, the PCR reaction also contained 2 pmoles each of the glucose-6-phosphate dehydrogenase-specific primers ZC3015 (Table 1; Sequence ID Number 12) and ZC3016 (Table 1; Sequence ID Number 13). After thirty cycles (one minute at 94°C, one minute at 60°C, ninety seconds at 72°C), the samples were phenol-chloroform extracted and 20% of each reaction was electrophoresed in agarose. The DNA was bidirectionally transferred to nitrocellulose membranes, and the filters were hybridized with either radiolabeled ZC3652, ZC3654, ZC3015 and ZC3016 (Sequence ID Nos. 14, 15, 12 and 13, respectively) or with the radiolabeled Eco RI-Xba I fragment of clone 45-A described above. Autoradiography of the hybridized blot showed that GlugR transcript was mainly confined to total rat brain and rat cerebellum; however, longer exposures showed a GlugR-specific transcript in both retina and testis.

Total RNA was prepared, as described above, from specific rat brain regions including frontal cortex, cerebellum, hippocampus, cortex, striatum, pons medulla, and the remainder of the brain. Single-stranded cDNA was synthesized as described previously using 20 μ g of total RNA in 50 μ l using Superscript reverse transcriptase (BRL) under conditions described by the manufactur r. After a one hour incubation at 42°C, the sampl s w re treated with RNAse (Boehringer Mannheim Biochemicals, Indianapolis, IN), phenol-chloroform extracted, and

10

15

20

25

30

35

ethanol precipitated. The samples were resuspended in wat r and half of each sample was subjected to PCR amplification. Each PCR amplification contained 40 pmoles of each of the GlugR-specific primers ZC3652 and ZC3654 described above (Sequence ID Numbers 14 and 15), 2 pmoles of each of the glucose-6-phosphate dehydrogenase-specific primers ZC3015 and ZC3016 (Sequence ID Nos. 12 and 13) and 2.5 U Tag I polymerase (Perkin Elmer Cetus) and conditions described by the manufacturer. After 35 cycles (one minute at 94°C, one minute at 60°C, ninety seconds at 72°C), the samples were phenol-chloroform extracted, and 20% of each reaction was electrophoresed in agarose. The DNA was transferred to a nitrocellulose membrane, and the filter was hybridized with the radiolabeled Eco RI-Xba I fragment of clone 45-A described above. Autoradiography of the hybridized blots showed a broad distribution of the Gluck transcript throughout the brain, although the frontal cortex and cerebellum appear to be somewhat enriched.

Southern analysis of rat and human genomic DNA was carried out using the method essentially described by Blin et al. (Nuc. Acids Res. 3:2303 (1976), which is incorporated by reference herein). Briefly, rat and human genomic DNA was prepared from the rat cell line UMR 106 (ATCC CRL 1661) and a human hepatoma cell line (ATCC HTB 52), respectively. The genomic DNA was digested with either Eco RI or Pst I, and electrophoresed through agarose. The DNA was transferred to a nitrocellulose membrane, and the membrane was hybridized with a radiolabeled 1.6 kb Pst I fragment from clone 45-A. Autoradiography of the hybridized blot suggest that the human gene has a similar sequence to the rat Gluck sequence, the Gluck gene contains at least one intron, and that there are a small number of closely related genes.

Expression in Mammalian Cells

The ntire GlugR cDNA insert was removed from the pVEGT' cloning vector by digestion with Not I and Xba I.

15

20

25

30

35

The ends w re blunted with DNA polymerase I (Klenow fragment) and dNTPs, and were then ligated with Eco RI (Smart) linkers. After linker ligation, the insert with Eco RI ends in kinased and ligated to Eco RI-cut and capped Zem228 expression vector. Bacteria were transformed with the ligation reaction and clones were characterized by restriction analysis and partial sequencing (see Fig. 4).

Cultured mammalian cells, such as BHK 570 and BHK ts13 served as host cells for expression. Twenty five μg of CsCl-purified DNA was precipitated with calcium phosphate and added to tissue culture cells in a 150 mm plate. After 4 hours the cells were subjected to a glycerol shock and were then put into non-selective medium. In some cases it may be necessary to include an antagonist to the Gluck in the medium to prevent expression of a cytotoxic response in those cells where the Gluck is expressed at levels high enough to cause a certain amount of autoactivation. Transiently expressed Gluck ligand binding activity or PLC activation, cells are harvested after 48 hours. Stable expression was detected after 2 weeks of selection. The Zem228 expression vector includes a promoter capable of directing the transcription of the Gluck gene, and a selectable marker for the bacterial neomycin resistance gene. Resistance to the drug G-418, an inhibitor of protein synthesis, was used to identify stably transfected clones. Presence of the SV40 ori region on the vector allows the expression construction to also be used for transient expression. In some instances it was preferable to include DNA for another selectable marker, the DHFR gene, in the transfection protocol. Selection with both G-418 and methotrexate allowed isolation of clones whose expression of GlugR can be subsequently amplified by the addition of increasingly higher concentrations of methotrexate to the culture medium.

Transfect d c ll lines $xpressing Glu_GR w r$ identified by the binding of 3H -glutamate t membrane

15

20

25

30

preparations from transf ct d cells. Cell lin s expressing low to moderate levels of GlugR ar used to set up functional screening assays.

Clones of BHK 570 and BHK TK ts13 cells expressing the rat G protein-coupled glutamate receptor cDNA were plated in two or three 150 mm maxi-plates culture dishes and were grown to confluency. The cells from each plate were scraped in 5 ml of PBS (phosphate buffered saline, Sigma Chemical Co., St. Louis, MO), which was was pre-chilled to 4 °C. The cells were removed to a chilled centrifuged tube, and the plates were each rinsed with 5 ml of chilled PBS and pooled with the cells. The chilled tubes were spun at 1,000 rpm for two minutes, and the supernatant was discarded. The cells were frozen at either -70°C or on dry ice. In some cases, the cells were left overnight at -70°C. The cells were thawed on ice and were resuspended in 10 ml of a buffer containing 30 mM Tris, pH 7.0, 2.5 mM CaCl, 1 mM PMSF, which was pre-chilled to 4°C, by homogenizing the cells for about 15 seconds. The suspension was poured into chilled centrifuge tubes. The homogenizer was rinsed with 10 ml of the same chilled solution, and the rinse was combined with the suspension. The centrifuge tubes were spun for fifteen minutes at 40,000 x g at 4°C, and the supernatant was discarded. The pellet was homogenized with a buffer containing 30 mM Tris, pH 7.0, 2.5 mM CaCl2, which was pre-chilled to 4°C. The homogenizer was rinsed with the chilled buffer, and the rinse was combined with the homogenate. The homogenate was spun as described above. The second homogenization was repeated on the resulting pellet. The final pellet was resuspended in between two and five milliliters of 30 mM Tris, pH 7.0, 2.5 mM CaCl2, which was pre-chilled to 4°C. Triplicate samples were prepared for each plus and minus quisqualate assay point such that 250 μ l aliquots of each homogenate sampl were added to the wells of a 96-well microtiter plate. buffer containing 30 mM Tris, pH 7.0, 2.5 mM CaCl2, which was pre-chilled to 4°C, a final concentration of 10 nM

5.

10

15

20

25

3.0

35

tritiated glutamic acid was added, and the soluti n was split in half. To one half, quisqualate was added to a final concentration of 1 mM. Two hundred and fifty microliter aliquots of either 30 mM Tris, pH 7.0, 2.5 mM CaCl₂, 5 nM tritiated glutamic acid and 500 mM quisqualate, or 30 mM Tris, pH 7.0, 2.5 mM CaCl₂, 5 nM tritiated glutamic acid were added to the triplicate samples. The samples were incubated for thirty minutes at room temperature. The samples were harvested onto glass filters and were immediately washed with ice-cold 30 mM Tris, pH 7.0, 2.5 mM CaCl₂ under vacuum using an LKB 1295-001 automated cell harvester (Pharmacia LKB, Piscataway, NJ). The filters were dried in a microwave oven and counted in a gamma counter.

Protein determinations were carried out using a Coomassie Blue-based assay from Pierce Chemical Company (Rockford, IL) under conditions set forth by the manufacturer. One hundred microliters of undiluted cell homogenate or BSA standard was added to 2 ml of reagent and the optical density was measured at 595 nm. Protein concentrations of the samples were taken from a standard curve generated using the BSA standards diluted in 30 mM Tris, pH 7.0, 2.5 mM CaCl₂.

The results of these assays showed that quisqualate was able to competitively bind the glutamate receptor expressed by the transfected BHK cells.

Functional screening of agonists and antagonists

BHK 570 cells expressing GluGR or mock-transfected BHK 570 cells are plated into 24-well tissue culture dishes at about 100,000 cells per well. After 24 hours, the cells are labeled with 0.2 μ Ci of myo-(2- 3 H) inositol (specific activity - 20 Ci/mmol; New England Nuclear,) per well. At the end of a 24 to 48 hour incubation, the cells are wash d with pr warmed DMEM (Dulb cco's Modified Eagles Medium; Product No. 51-432, JRH Biosciences, Lenexa, KS) which has been buffered to pH 7.4 with H pes

buffer (Sigma Chemical Co.) containing 10 mM LiCl, and are incubat d for five minutes at 37. The selected drugs are then added and the cells are incubated for an additional thirty minutes at 37°C. The reaction is stopped by placing the cells on ice, and the cells are lysed by aspirating off the media and adding 0.5 ml of cold DMEM and 0.5 ml of ice-cold 10% perchloric acid. After ten minutes the cell lysate is transferred to a tube on ice containing 250 μ l 10 mM EDTA, pH 7.0. samples are neutralized with 325 μ l of 1.5 M KOH in 60 mM Hepes Buffer. After the precipitates settles, 1.0 ml of the supernatant is applied to an Amprep minicolumn (Amersham, Arlington Heights, IL, RPN1908). phosphates are eluted off the column and samples are counted in a scintillation counter. A positive response is indicated by an increase in labeled inositol phosphate levels.

EXAMPLE II

20

25

30

35

15

10

Screening for additional glutamate receptor subtypes

Additional glutamate receptor subtypes were isolated using probes derived from clone 45-A. Glutamate receptor subtypes were isolated from a total rat brain cDNA library in Lambda Zap II, which was size selected for inserts of 3 kb before ligation (prepared for Terry Snutch, Ph.D., University of British Columbia, Vancouver, British Columbia, Canada by Stratagene Cloning Systems, La Jolla, CA) and a rat cerebellum cDNA library in Lambda Zap II, which was size selected for inserts of 3 kb before ligation (Stratagene Cloning Systems, La Jolla, CA).

The total rat brain library and the rat cerebellum library were plated out with <u>E. coli</u> XL-1 cells onto NZY agar plat s (Table 3) to obtain approximately 2.1 x 10⁶ plaques. Clon 45-A, encoding subtype la, was digested with Pst I to is late the 1.3 and 1.6 kb fragm nts. The 45-A Pst I fragments were labeled by random priming using

the Am rsham rand m-priming kit (Amersham, Arlington Hts, IL). Duplicate lifts wer pr pared from the plates, and the filters were hybridized with the probes in 50% formamide at 37°C. After an overnight hybridization, the filters were washed in 2x SSC + 0.1% SDS at 50°C. Positive plaques were isolated by several rounds of dilution plating and repeated screening with the random-primed probes.

10

15

Table 3

NZY Agar

To 950 ml of deionized water, add:

10 g NZ amine: Casein hydrolysate enzymatic (ICN Biochemicals)

5 g NaCl

5 g bacto-yeast extract

1 g casamino acids

2 g MgSO, 7H,0

20

Shake until the solutes have dissolved, Adjust to pH 7.0 with 5 N NaOH (approximately 0.2 ml). Adjust the volume of the solution to 1 liter with deionized $\rm H_2O$. Sterilize by autoclaving for 20 minutes.

25

20x SSC

Dissolve 175.3 g NaCl and 88.2 g sodium citrate in 800 ml H₂O. Adjust the pH to 7.0 with a few drops of 10 N NaOH. Adjust the volume to 1 liter with H₂O. Sterilize by autoclaving.

30

35

Plasmid DNA was prepared from positive plaques using the Bluescript system (Stratagene Cloning Systems). The plasmid DNA was subjected to restriction analysis and Southern blot analysis (Sambrook et al., ibid., which is incorporat d h rein by reference). Tw clones, SN23, derived from the total rat brain library, and SR2, derived from the rat cerebellum library, were id ntified

10

15

20

25

30

35

as b ing differ nt than the 45-A clone and were sequenc d. Sequenc analysis showed that they represented two additional subtypes. SN23 encodes subtype 1b, which contains an additional 85 bp exon that encodes a new stretch of 20 amino acids and a stop codon in the intracellular domain, is 292 amino acids shorter than the 45-A clone. The nucleotide sequence and deduced amino acid sequence of clone SN23 are shown in Fig. 7. SR2 was found to contain a partial cDNA sequence encoding subtype 2a, which is a novel sequence that shares a 42% homology to the transmembrane domains and extracellular domain of the 45-A clone.

A complete subtype 2a clone was obtained by rescreening both libraries as described above with the radiolabeled 1.3 kb Pst I fragment from clone 45-A and a radiolabeled 1.4 kb Eco RI-Pvu II fragment from SR2. additional clones were obtained. SN30, derived from the total rat brain library, contained the entire subtype 2a coding sequence. The nucleotide sequence and deduced amino acid sequence of clone SN30 are shown in Fig. 8. SR13, derived from the rat cerebellum library, contained an incomplete sequence of a new receptor subtype, 2b. Sequence analysis of SR13 showed that the coding sequence was incomplete at the 3' end and was virtually identical to the SN30 sequence except that it contained a 610 base pair deletion within the 3' terminus of SN30. sequence of the cDNA insert in clone SR13 is shown in Figure 9.

The complete 3' end of the subtype 2a clone was generated using PCR amplification and an oligonucleotide containing a sequence unique to SR13 (ZC4520, Table 4) and an oligonucleotide corresponding to a sequence near the 3' end of the 3' non-translated region of SN30 (ZC4519, Table 4). DNA was prepared from plate lysates of the original plating of ach library. Each plate produced a pool of clones. For the PCR reactions, ten nanograms from each library and 100 pmol of each oligonucleotide were combined in a reaction volume of 50

μl containing 50 mM KCl, 10 mM Tris-HCl pH 9.0, 1.5 mM MgCl₂, 0.1% Triton X-100, 0.01% gelatin, 0.2 mM each deoxynucleotide triphosphate and 2.5 units of <u>Thermus aquaticus</u> (Taq) DNA polymerase (Promega Corporation, Madison, WI). The reaction mixture was overlaid with mineral oil. After five cycles (30 seconds at 94°C, 30 seconds at 45°C, 1 minute at 50°C) and twenty-five cycles (30 seconds at 94°C, 30 seconds at 45°C, 1 minute at 72°C) the amplified DNA was removed for analysis.

10

15

20

25

30

35

Table 4

Degenerate Oligonucleotide Primer Sequences (5' - 3')

ZC4519

TTT ATT AGA AAT GTT CTC GGT

ZC4520

CCT CTT CCA TAT TTT TCC ATT

ZC4559

ATA AGA ATT CAT NKR YTT NGC YTC RTT RAA

ZC4560

ATA AGA ATT CTT YRA YGA RAA NGG NGA YGC

ZC4561

ATA AGA ATT CGC NGG NAT HTT YYT NKG NTA

ZC4562

ATA AGA ATT CTA NCM NAR RAA DAT NCC NGC

ZC4563

ATA AGA AAT CAN GTN GTR TAC ATN GTR AA

An aliquot from each reaction was electrophoresed on agarose and transferred to nitrocellulose for Southern analysis. Southern analysis of the PCR products showed that a 460 bp fragment corresponding to the 3' end of the 2b sequence was present in several pools. One of the pools that produced the correct size PCR product encoding the 3' sequence of the 2b subtype was diluted and scre ned with radiolabeled ZC4519 and ZC4520 (Table 4). Phag that hybridize to both radiolabeled ZC4519 and ZC4520 are picked, eluted, dilut d, plated and rescreen d with the oligonucleotide prob s. The screening is

15

20

25

30

35

repeated until a pure clone is obtained. The pure clone is sequenced, and a full-length clone is constructed using the most convenient restriction enzyme(s).

Based on an alignment of the deduced amino acid sequences of subtypes 1a and 2a, strategies were designed for cloning additional subtypes using PCR amplification. Degenerate oligonucleotide families were prepared to encode conserved amino acid sequences in the sixth transmembrane domain, a region surrounding the conserved amino acid sequence Phe-Asp-Glu-Lys, the third cytoplasmic loop, and the second transmembrane domain (Table 4).

Glutamate receptor cDNA sequences were amplified with pairs of degenerate primers from Table 4 using the PCR method on cDNA from the total rat brain library, the cDNA from the rat cerebellum library, a rat cortex cDNA library or a rat hippocampus cDNA library (both obtained from Michael Brownstein, National Institutes of Health, Bethesda, MD). The primers also each contained a 5' tail of 10 nucleotides, which provided convenient restriction enzyme sites. For each PCR reaction, ten nanograms from the library and 100 pmol of the oligonucleotide pools ZC4563 and ZC4560 (Table 4) were combined in a reaction volume of 50 μ l containing 50 mM KCl, 10 mM Tris-HCl pH 9.0, 1.5 mM MgCl2, 0.1% Triton X-100, 0.01% gelatin, 0.2 mM each deoxynucleotide triphosphate and 2.5 units of Taq DNA polymerase. The reaction mixture was overlaid with mineral oil. After five cycles (30 seconds at 94°C, 30 seconds at 45°C, 1 minute at 50°C) and twenty-five cycles (30 seconds at 94°C, 30 seconds at 45°C, 1 minute at 72°C) the amplified DNA was removed for analysis.

An aliquot from each reaction was electrophoresed on an agarose gel. Southern analysis of the gel was performed using essentially the method described by Sambrook et al. (ibid.) and random-primed fragments cov ring the entir coding regions from both the subtype la and 2a clones. The autoradiographs showed that the PCR reaction gen rat d fragments of novel siz that wer

WO 92/10583 PCT/US91/09422

61

differ nt from either the la or 2a subtyp . The PCR-generated fragments w r el ctrophoresed on an agarose gel. Regions corresponding to the unique-sized receptor-related products were excised and electrophoresed onto NA45 paper (Schleicher and Schuell, Keene, NH). The purified fragments were recovered using essentially the method described by the manufacturer, digested with Eco RI and ligated to plasmid pVEGT! that had been linearized by digestion with Eco RI and treated with phosphatase to prevent recircularization. The ligation mixtures were transformed into E. coli strain DH10b cells. Transformants were picked and replica plated onto nitrocellulose filters and screened using random-primed probes from the la and the 2a clones. Forty-eight colonies were picked for restriction analysis and sequencing.

10

15

20

25

30

35

DNA sequences from the cDNA from the total rat brain library and the cDNA from the rat cerebellum library were each amplified and analyzed using the methods described above and oligonucleotide ZC4559 in combination with either ZC4561 or ZC4559 (Table 4).

A rat cortex cDNA library and a rat hippocampus cDNA library (both obtained from Michael Brownstein, NIH) are subdivided into 30 pools of 10,000 colonies. Plasmid DNA is prepared from each pool, and the DNA is subjected to Southern analysis after restriction digestion of the pools with Bam HI and Xho I or by PCR amplification of each pool using the degenerate oligonucleotides of Table 4. The library pools containing DNA that hybridize to the probes and appear to contain a full-length cDNA are subdivided. The plasmid DNA is prepared and screened as described above. Positive pools are again divided and the procedure is continued until the pool is reduced to The clones are subjected to restriction pure clones. analysis and partial sequence analysis. Clones that represent distinct glutamate rec ptor homologs are completely s qu nced. Full length clones are g nerat d by subjecting the original pools to PCR amplification

using an oligonucleotide primer specific to the SP6 promot r at the 5' end of the cDNA insert and an antisense oligonucleotide primer corresponding to the 5' end of the most complete cDNA to identify pools that contain the longest glutamate receptor homolog cDNA. The pool is then diluted and rehybridized with the probes as described above to isolate a full length cDNA clone.

Expression of Glutamate Receptor Subtypes

.

15

10

5

20

25

30

Plasmid SN30, which comprises the subtype 2a cDNA, was digested with Eco RI to isolate the subtyp 2a cDNA.

The Eco RI fragment was ligated with Eco RI-lineariz d

Zem228R. A plasmid containing the insert in the correct ori ntation was digested with Bam HI t isolate the cDNA

Complementary DNA sequences encoding subtypes 1b and 2a were subcloned first into the mammalian expression vector Zem228R to obtain convenient terminal restriction The cDNAs were then subcloned into pVEGT'. sites. cDNA sequence encoding subtype 1b was constructed by replacing the 3' terminal portion of subtype la described in Example I with the analogous portion of subtype 1b from SN23. Plasmid SN23 was digested with Kpn I and Xba I to isolate the fragment containing the 3' terminus of the 1b subtype. The plasmid containing the subtype la coding sequence (45-A) in Zem228R was digested with Kpn I and Xba I to isolate the vector containing fragment. The vector containing fragment is ligated to the Kpn I-Xba I fragment from SN23. The resulting plasmid comprises the MT-1 promoter, the subtype 1b cDNA and the hGH This plasmid was transfected into the BHK terminator. 570 cell line essentially as described in Example I to obtain stably transfected cell lines expressing the subtype 1b receptor. The subtype 1b cDNA fragment was isolated as a Bam HI fragment, which was ligated with pVEGT' that had been linearized with Bam HI. A plasmid containing the cDNA sequence in the correct orientation was used to synthesize RNA in an in vitro system. RNA was injected into oocytes as described above.

sequence. The Bam HI fragment comprising the subtype 2a cDNA was ligated with Eco RI-linearized pVEGT'. A plasmid containing the cDNA in the correct orientation was used to synthesize RNA in an in vitro translation. The RNA was injected into frog oocytes as described above.

EXAMPLE III

Generation of antibodies to glutamate receptor subtypes

10

15

Receptor subtype-specific polyclonal antisera were generated in rabbits using standard immunization techniques. Synthetic peptides (Table 5) were designed from the cloned receptor sequences. The peptides were conjugated to keyhole limpet hemocyanin, and each antigen was used to immunize two animals. For each peptide, the animals were injected with $100-200~\mu g$ of conjugated peptide divided among three subcutaneous sites. The animals were immunized at three-week intervals and bled via an ear vein 10 days after the third and subsequent immunizations.

20

Table 5

25	Subtype	Seq. ID	Peptide <u>Sequence</u>	Apparent Location
25	1a	21	RDSLISIRDEKDGLNRC	extracellular
		22	DRLLRKLRERLPKARV	extracellular
(C. C.)		23	EEVWFDEKGDAPGRYD	extracellular
•		24	EFVYEREGNTEEDEL	cytoplasmic
30	* (- *	25	PERKCCEIREQYGIQRV	extracellular
· · · · · · · · · · · · · · · · · · ·		26	IGPGSSSVAIQVQNLL	extracellular
•		27	IAYSATSIDLSDKTL	extracellular
1. *	1b	28	KKPGAGNAKKRQPEFS	cytoplasmic
, ·		29	PEFSPSSQCPSAHAQL	cytoplasmic
.35	2a	30	DKIIKRLLETSNARG	extracellular
* *		31	VNFSGIAGNPVTFNEN	extracellular
		32	GEAKSELCENLETPAL	cytoplasmic
	2b	33	PARLALPANDTEFSAWV	cytoplasmic
	,	•	· •	

15

30

35

40

Anti-peptide antibodies w re purifi d by affinity purification using the ProtonTM Kit (Multiple Peptide Systems (San Diego, CA). Purified antibodies were stored in column elution buffer and neutralizing buffer (supplied by Multiple Peptide Systems). Bovine serum albumin was added to a concentration of 1 mg/ml, and sodium azide was added to a concentration of 0.05%. The antibodies were stored at 4°C or in small aliquots at -20°C.

Antibodies generated from the peptides listed in Table 6 were used to detect G protein-coupled glutamate receptors by Western blot analysis of membranes prepared from transfected cell lines that were stably expressing the subtype 1a or subtype 1b receptors. Control cell lines were transfected with vector alone.

Table 6

Analysis of Antibodies Raised to Peptides

20	Antibodies to Peptide Sequence	Seq. ID	Location	Western
*	RDSLISIRDEKDGLNRC	21	extracellular	+++ with bkgd
	DRLLRKLRERLPKARV	22	extracellular	+
25	EEVWFDEKGDAPGRYD	. 23	extracellular	++++ low bkgd
	EFVYEREGNTEEDEL	24	cytoplasmic	++++ low bkgd
	KKPGAGNAKKRQPEFS	28	cytoplasmic	+ for la - for lb
	PEFSPSSQCPSAHAQL	29	cytoplasmic	+++ for 1b low bkgd

Transfectants that were stably expressing either the la or 1b subtype were each grown to confluency in five to ten 150 mm plates. Each plate was first washed twice with 15 ml of cold PBS and then 20 ml of ice cold 10 mM NaHCO3 was added to each plate. The cells from each plate were scraped off the plates with a rubb r spatula and transf rr d to a glass dounc homog niz r n ice. Th cells were disrupted with ten strokes of the B pestle. Th homog nates from each plate wer c mbined

and centrifuged for thirty minutes at 3000 rpm at 4°C. The pellets were resuspended in 4-8 ml of 10 mM NaHCO3 using a 22 g needle and syringe, and 69% sucrose was added (6-12 ml) to each sample until an index of refraction of 1.410 was reached. The samples were transferred to a high speed centrifugation tube, and each sample was overlayed with 42% sucrose. The samples were centrifuged for two hours at 25,000 rpm at 4°C. samples were collected by gently floating the membranes off the 42% sucrose layer by adding 1 ml of 10 mM NaHCO, and resuspending the membranes by carefully stirring the upper layer. The upper layer was transferred to a fresh tube on ice. The purified membranes were centrifuged at 10,000 rpm at 4°C and the pellets resuspended in 10 mM NaHCO. The purified membranes were then adjusted to a final protein concentration of 1-2 µg/ml.

Ten to twenty micrograms of each purified membrane preparations were diluted with 2x SDS-mercaptoethanol buffer (100 mM Tris HCl (pH 6.8), 200 mM dithiothreitol, 4% SDS, 0.2% bromophenol blue, 20% glycerol). samples were incubated for 15 minutes at 37°C followed by boiling for 5 minutes. The samples were subjected to SDS-PAGE on 4-15% gradient gel. The samples were electrotransferred to nitrocellulose using the method essentially described by Towbin (Proc. Natl. Acad. Sci. <u>USA</u> 76: 4350-4354, 1979; which is incorporated herein by reference in its entirety). After transfer, the nitrocellulose was cut into strips such that each strip contained a control and receptor samples. nitrocellulose was preincubated in blocking buffer and then incubated with a dilution of either the preimmune serum or the serum collected after antigenic stimulation (serum from later bleeds (i.e. those after four antigen stimulations) were diluted 1:1500). After washing, a horse radish p roxidase-conjugat d goat anti-rabbit antibody (Bi -Rad Laboratories, Richmond, CA) diluted 1:2,500 was added and after incubation and washing, th h rse radish peroxidase substrate (Bio-Rad Laboratories)

15

10

20

25.

30

WQ 92/10583 PCT/US91/09422

66

was added and the color reaction was initiated. The reaction was stopped by rinsing the filters in distilled water. Table 6 shows the results of the Western blot analysis.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Mulvihill, Eileen R. Hagen, Frederick S. Houamed, Khaled M. Almers, Wolfhard
- (ii) TITLE OF INVENTION: G PROTEIN-COUPLED GLUTAMATE RECEPTORS
- (iii) NUMBER OF SEQUENCES: 33
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Townsend and Townsend
 - (B) STREET: One Market Plaza, Steuart Street Tower
 - (C) CITY: San Francisco
 - (D) STATE: California
 - (E) COUNTRY: USA
 - (F) ZIP: 94105-1492
 - (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 07/672,007
 - (B) FILING DATE: 18-MAR-1991
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 07/648,481
 - (B) FILING DATE: 30-JAN-1991
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 07/626,806
 - (B) FILING DATE: 12-DEC-1990
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Parmelee, Steven W.
 - (B) REGISTRATION NUMBER: 31,990
 - (C) REFERENCE/DOCKET NUMBER: 13952-6PC
 - (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (206) 467-9600
 - (B) TELEFAX: (415) 543-5043
- (2) INFORMATION FOR SEQ ID NO:1:

-	(i)	(A) (B)	ENCE LEN TYP STR TOP	GTH: E: n ANDE	430 ucle DNES	ic a S: s	se p cid ingl	GILL				er.				٠	
		TMME	CULE	E SC	URCE	:					nc.	, ,		, .			*
		(A) (B)	TURE: NAM LOC	E/KI LATIO	ON: 3	377.			,		*				-		· · · · · · · · · · · · · · · · · · ·
	'vi)	SEQ	JENCI	E DES	CRI	PTIO	N: S	EQ I	סא ס	:1:					•		40
CCGAG	2220	GG C	TGCA	TCC	r cto	BACC:	rgag	ACC	AATA	GCT (GTGT(CTAC	CC G	GACT(CAGCG		60
CCGAC			6001	OTT N A	c eco	acce.	CGCA	TTG	GACA	CCT (GATC	CACA	CA C	CTTC	GGGCA	٠	120
TCCAC	:CTC	AC C	GULA	-Tww.			-	336	aà cc	ccc (CCAG	GGTG'	rg g	GAGC	GGTCG		180
CCAG	IGAA	AA A	CCGČ	GACT	r GA	T-T-T-T-	CIGG	MAG			3 CMC	CACC	CA G	AGAA.	GGTCG AGCGT	· · · ·	240
TGGA	GAC	CA G	CAGG	AGGA	A GC	GGAG	GGGA	GAG	GGGC.	AGT 4	AGTG	GAGG			AGCGT	, 1	300
TGAA	CCAG	CT G	TGTT	GGCC	G AA	GGCA	CGAA	ACG	GCAA	AAG	GCAG	CGGT	GAG :	CATC	TGTGT		
COMM		: حصات رو	GGAA	CCTG	C AG	GCAG	GACC	GGC	GTGG	GAA	CGTG	GCTG	GC C	CGCG	GTGGA	•	360
CCGC	GTCI	TC G	CCAC				شد، ب			G AT	TT T	CTT	C CC	A AT	G		409
ATC	TTT Phe	TTG Leu	GAG Glu 15	ATG Met	TCC	ATT Ile	TTG Leu	CCC Pro 20	agg Arg	ATG Met	CCT Pro	GAC Asp	AGA Arg 25	aaa Lys	GTA Val		457
Leu	Leu	Ala 30	GGT Gly-	ATG	per	Jer	35					40					505
Asp	Val	Ile	ATC Ile	GTY	ALG	.50					55						553
Glu 60	Lys	Val	CCC Pro	GIU	65	23.0	-2-	*		70					, /5		601
Ile	Gln	Arg	GTG Val	80	Wīd	Mec			85	-				90			649
GAC Asd	CCG	GTG Val	CTC Leu	CTG Leu	CCC Pro	AAC Asn	ATC	ACT	CTG Leu	GGC Gly	AGT Ser	GAG Glu	ATC	CGG	GAC Asp		697

		-,									•	1.1						
	•	~ ,		95					100			•		105				
	TCC Ser	TGC Cys	TGG Trp	CAC His	TCT Ser	TCA (var .	VIG	Len	G AA Glu	CAG Gln	AGC Ser	ATC Ile 120	GAA Glu	TTC Phe	ATC Ile	*	745
			110					113					-, -					
	AGA Arg	GAC Asp	TCC	CTG Leu	ATT	TCC . Ser	ATC Ile	CGA Arg	GAT Asp	GAG Glu	AAG Lys	unb	GGG	Leu	AAC Asn	CGA Arg		793
		125					T30	٠.		•								
	TGC	CTG	CCT	GAT	GGC	CAG	ACC	CTG	CCC	CCT	GGC	AGG	ACT	AAG	AAG	CCT		841
	Cys 140	Leu	Pro	Asp	Gly	Gln 145	Tnr	Leu	Pro	PLO	150	nry	*	-3 -		155		
	ATT	GCT	GGA	GTG	ATC	GGC	CCT	GGC	TCC	AGC	TCT	CTG	GCC	ATT	CAA	GTC Val		889
	Ile	Ala	Gly	Val	11e	GLY	Pro	Gly	Ser	Ser 165	Ser	Val	Ala	TIE	170	AGT	. :	•
	CAG	AAT	CTT	CTC	CAG	CTG	TTC	GAC	ATC	CCA	CAG	ATC	GCC	TAT	TCT	GCC		937
	Gln	Asn	Leu	Leu 175	Gln	Leu	Phe	Asp	Ile 180	Pro	Gln	Ile	ALA	185	Ser	Ala	*	
	àCA	AGC	ATA	GAC	CTG	AGT	GAC	AAA	ACT	TTG	TAC	AAA	TAC	TTC	CTG	AGG		985
	Thr	Ser	Ile	Asp	Leu	Ser	Asp	Lys 195	THE	Leu	Tyr	Lys	TYT 200	Pne	Leu	Arg		v
8			190			• .•					000) MC	- CERC	GAC	`` `` ልጥል	GTC		1033
	GTG	GTC	CCI	TCT	GAC	ACT	TTG	CAG Gln	GCA Ala	AGG	Ala	Met	Leu	Asp	Ile	Val		
	AT	205		, 361	, Aup		210			•	-	215		٠.	• .		0	•
	330	CCT	ጥልር		TGG	ACC	TAT	GTC	TCA	GCA	GTC	CAC	ACA	GAA	GGG	AAT		1081
	Lys 220	Arg	Tyr	Asn	Trp	Thr 225	Tyr	Val	Ser	Ala	Val 230		Thr	Glu	GTÅ	Asn 235		
	ma c	GGC	GAC	. AGT	GGA	ATG	GAT	GCT	TTC	AAA	GAA	CTG	GCT	GCC	CAG	GAA		1129
	Tyr	Gly	Glu	Ser	Gly 240	Met	Asp	Ala	Phe	Lys 245	914	Leu	Ala	Ala	250			
	ccc	- COLO	ነ ጥርረ	- <u>አ</u> ጥር	C GCA	CAC	TCG	GAC	: AAA	ATC	TAC	: AGC	AAT	GCT	GGC	GAG Glu	•	1177
	Gly	Leu	Cy	3 Ile 25!	s Ala	His	Ser	Asp	260	TTE	TYE	Ser	. Asi	265		Glu		
	120	. 200	- Made	ም ር እር	c ccc	CTC	CTG	: CG7	. AA	CTC	: CGG	GAG	CGG	CII	CCC	AAG Lys		1225
	Lys	Set	Pho 27	e asi	o Arc	Leu	Leu	275	i nas	Leu	l Arc	, Glu	280		Pro	Lys	*	+
		8		· _	· 		mme.	• TC	r GAC	: GG(. ATC	: AC	A GTO	G CGC	GGG	TTA Leu	***	1273
	GCC	AGG	GT Va	r Gr 1 Va	G.GT(1 Va)	L Cys	Phe	Cys	s Glu	. Gly	Met			L Arg	GI	Leu		
	•	285	5	•		•	290	,				•						
	CTC	G AG	r GC	C AT	G CG(C CGC	CTC	GG	C GT	C GT	GGG	GA	G TT	CTC	CT	C ATT		1321
	Let 300	ı Şei	r Al	a Me	t Arg	305	. Ter	T GT	y va.	L Va.	310	0				315		•
,	cci	A AC	r GA	T GG	A TG	G GCI	GAG	C AG	A GA	r GA	A GT	C AT	C GA	A GG	C TA	r GAG r Glu		1369
	Gly	y Se	r As	p Gl	y Tr	b vrs	Ası	o Ar	g As	p Gl: 32	<u> </u>	1 II	GT,	n GT	33	r Glu O	_	
					J B 1	_						٠.		•		•		

SUBSTITUTE SHEET

	•				•	-	. ,			70	- 00	,			•			•
	GTC Val	GAA Glu	GCC Ala	AAC Asn	GGA Gly	GGG Gly	ATC Ile	ACA Thr	110	AAG Lys	CTT Leu	CAG Gln	TCT Ser	CCA Pro	GAG Glu	GTC Val	o X o	1417
	Y W.Z			335					340	•				343	•	•	5	
	AGG	TCA	TTT	GAT	GAC	TAC	TTC	CTG Leu	AAG Lys	CTG Leu	AGG Arg	CTG Leu	GAC Asp	ACC Thr	AAC Asn	ACA Thr	*	1465
	Arg	Ser	350	Asp	vah	-1-		355	-		. –		360	,				•
		• •	•			e			mcc	CAA	CAT!	CGC	TTC	CAG	TGT	CGC		1513
	AGG	AAT	CCT	TGG	TTC	CCT	GAG	Phe	Trp	Gln	CAT His	Arg	Phe	Gln	Cys	Arg		
	Arg	Asn 365	Pro	Trp	Pne	FLO	3.70					375						
						mma	~ \ \ \ \		CCC	AAC	TTT Phe	AAG	AAA	GTG	TGC	ACA		1561
	CTA	CCT	GGA	CAC	CTC	Leu	GAN	Asn	Pro	Asn	Phe	Lys	Lys	Val	Cys	Thr	,	
	380	Pro	GTÄ	UTO	5 54	385					390		·	,	. :	395	V	r*
	i.	• :	6 C.	, · _				***	ጥእጥ	GTC	CAG	GAC	AGC	AAA	ATG	GGA Gly	,	1609
	GGA	AAT	GAA	AGC	TTG	GAA	GAA	Asn	Tyr	Val	Gln	Asp	Ser	Lys	Met	Gly		
•	LLA	F JN	GIU	261	400	324	* .		1	405		•		-	410			4
		`.*				•	·			CON	CATT	GGG	CTG	CAG	AAC	ATG		1657
	TTT	GTC	ATC	AAT	GCC	ATC	TAT	GCC	ATG	Ala	His	Gly	Leu	Gln	Asn	ATG Met		
•	Phe	Val	Ile	ASN 415	ATa	TIE	TÄT	·VTA	:420	•••	-			425			•	•
					1		*	·			conc.	mem	ĊĀŦĖ	CCT	ATG	AAA		1705
	CAC	CAT	ĢCT	CTG	TGT	CCC	GGC	CAT	GIG Val	GGC Glv	Leu	CVS	ASD	Ala	Met	aaa Lys		
:	His	His	Ala 430	Leu	cys	PIO	GTA	435	744	4-1 ,		* *	440				8	
		• ,					,				~~~	3 M/C	222	TCC	тст	TŤT	•	1753
	CCC	ATT	GAT	GGC	AGG	AAG	CTC	CTG	GAT	Phe	Teu	Ile	Lvs	Ser	Ser	TTT		
	Pro	Ile	Asp	Gly	Arg	ГÃЗ	450	Leu	, vsh	2110		455	4	14				
	: '	445				•						CNC		ccc	GAT	GCT	• •	1801
	GTC	GGA	GTG	TCT	GGA	GAG	GAG	GTG	TGG	TTC	GAT	GAG	LVS	Gly	Asp	GCT	*100	
	Val	Gly	Val	Ser	Gly	465	GIU	Val	, iip	2110	470		, - -			Ala 475		
	460	N.T	•			. •		•		· .			633	COTT		· · CGC		1849
	CCC	GGA	AGG	TAT	GAC	ATT	ATG	IAA :	CIG	CAG	TAC	ACA The	GAA	Ala	ASI	CGC	-ic-	
	Pro	Gly	Arg	Tyr	Asp) Tīg	Met	: ASI	i Ten	485		1			490	Arg		-
٠.			•	* *	480		- %-	: '		-				-	አአማ	እጥጥ ·		1897
	TAT	GAC	TAT	GTC	· CAC	GTC	GGG	ACC	TGG	CAT	GAA	. GGA	GIG Val	Cru Lev	ASI	ATT		
	Tyr	Asp	Tyr	. Agi	. Hls	Val	Gly	Tni	500		GIU	GLY	761	505	,	lle		
				495)	· ;		•		٠.,								1945
	САТ	GAT	TAC	: AAA	ATC	CAG	ATC	AA	CAAA	AGC	GGA	ATG	GTA	CGA	Set	GTG Val	: -	1,740
-1	Asp	Asp	Туз	Lys	Ile	e Glr	1 Met	י אפי	. .	; Ser	GTA	Met	. va. 520)	, 50.	val		0
		*	510)					•								1	1003
,		3 CF	n cac	ء دريا	r TGC	TT	AAG	GG'	r CAC	TTA :	AAG	GTC	AT	CGG	AA	A GGA B Gly		1993
	CVS	. AGI	Glu	Pro	Cys	Let	r na:	S GE	y Gli	ı:Ile	Lys	Val	: : Ile	a Arg	i nă:	s Gly	* -	14 1
		525	5		٠		231			. •								
•				- MC(። ግ ጥርር	TGO	TA	C TG	C ACC	GCC	TGC	AA	A GAG	3 AAT	r GA	G TTT	•	2041
	GAA	i Val	i AG L Sei	CVS	Cy:	s Tri	o Ile	в Су	s Thi	r Ale	Cys	Lys	s Gli	ı Ası	J GT	2 Phe 555	•	•
	540) [34:	9											
				7 634	े संस्कृत -	· C AC	C TG	C AG	A GC	C TG	r gac	CTC	G GG	G TG	g ŢG	G CCC		2089
	GTC	CAC	AU t	, was	.	- 			an y ages y	175	SH	EE	Γ					•
	. •						SU	B 5	1111	مطا ۾ ز	SH						*	

Val. Gin Asp Glu Phe Thr Cys Arg Ala Cys Asp Leu Gly Trp Trp Pro 565 AAC GCA GAC CTC ACA GGC TGT GAG CCC ATT CCT GTC CGT TAT CTT GAG Asn Ala Glu Leu Thr Gly Cys Glu Pro Ile Pro Val Arg Tyr Leu Glu 575 2137 TGG AGT GAC ATA GAA TCT ATC ATA GCA CTC GTT CTT CTC TTC GGC ATT TCT TGC CTG GGC TTP Ser Asp Ile Glu Ser Ile Ile Ala Ile Ala Phe Ser Cys Leu Gly 590 2185 ATC CTC GTG AGC CTG TTT GTC ACC CTC ATC TTC GTT CTG TAC CGG GAC IILe Lau Val Thr Leu Phe Val Thr Leu Ile Phe Val Leu Tyr Arg Asp 605 2233 ACA CCC GTG GTC AAA TCC TC AGT AGG GAC CTC TGC TAT ATC ATT CTG TAC CTG GGT ATT TC CTC GGC ATC TTC GTG TAC CTG TAT ATC ATT CTG TAC CGG GAC GGT ATT TC CTC GGC ATT TTC GTG CAC CTC ATC TCT ACC GG AAA ACC ATC GGT ATT TC CTC GGC GGT ATT TTC CTC GGC ATT TTC GTG CAC CTC ACT ACC ACA ACC ACA TCC TC TAC GTC CTC ACC CTC ACT ACC ACA TCC TC CAC TC CTC ACC CTC ACT ACC ACA TCC TC CAC TC CTC ACC CTC TAC TTT GTG TAC CTC TAC TTC TTT TTT TTT Ser Cys Tyr Leu Gln Arg Leu Leu Val Gly Leu Ser 665 2377 TCT GCC ATG TGC TAC TCT GGT TTA GTG ACC AAA ACC AAT GGT ATT GCA TCT GGT TTA GTC ACC TC TATT GGA ACC ATC TCT GGT TTA GTC ACC CTC ATT CTT TA GTC ACC ACC ACC ATC TCT GGT TTA GTC ACC ACA ACC ATC TCT GGT TTA GTC ACC ACA ACC ATC TCT GGT TTA GTC ACC CTC ATT TTA GGA ACT ATC TTT AGT ILeu Val Thr Lys Thr Asn Arg Ile Ala GTC TCC TCT GGT TTA GTC ACC ACC ACA ACC ATT TC GGA GGC ATC TTG GTC TGC TTT AGT ACC TC ATT CTT ATT AGT ACT TCT GGT TTA GTC ACC CTC ATT TTT AGT ACC TCT ATC ATC ATG GGC CC ATC TTT ATT AGT ACT TTT AGT ACT TTT AGT ACC TTT ATT AGT TTT AGT ACC TTT ATT AGT ACT ATT AGT TTT AGT ACC TTT ATT AGT ACT ATT AGT TTT AGT ACC TTT ATT AGT ACT ATC ATC ATC ATC ATC ATC ATC ATC AT		•												•					•
ASIN Ala Glu Leu Thr Gly Cys Glu PFO 1880 TOS 575 TOG AGT GAC ATA GAA TCT ATC ATA GCC ATC GCC TTT TCT TGC CTG GGC TTP Ser Asp Ile Glu Ser Ile Ile Ala Ile Ala Phe Ser Cys Leu Gly 590 ATC CTC GTG ACG CTG TTT GTC ACC CTC ATC TTC GTT CTG TAC CGG GAC ILE Leu Val Thr Leu Phe Val Thr Leu Ile Phe Val Leu Tyr Arg Asp 605 ACA CCC GTG GTC AAA TCC TCC AGT AGG GAG CTT GC TAT ATC ATT CTG TAT PTO Val Val Lys Ser Ser Ser Arg Glu Leu Cys Tyr Ile Ile Leu 622 GCT GGT ATT TTC CTC GGC TAT GTG TGC CTT TTC ACC CTC ATC GCT AAA A 2329 ALA Gly Ile Phe Leu Gly Tyr Val Cys PTO PHe Thr Leu Ile Ala Lys 640 CCT ACT ACC ACA TCC TGC TAC CTC CAG CGC CTC CTA GTT GCC CTT CT PTO Thr Thr Thr Ser Cys Tyr Leu Gln Arg Lau Leu Val Gly Leu Ser 655 TCT GCC ATG TGC TAC TCT GGT TTA GTG ACC AAA ACC AAT CGT ATT GCA Ser Ala Met Cys Tyr Ser Ala Leu Val Thr Lys Thr Asn Arg Ile Ala 665 CGC ATC CTG GCT GGC AGC AAG AAG AAG AAG ATC TGC ACC CGG AAG CCC AGA Arg Ile Leu Ala Gly Ser Lys Lys Lys Ile Cys Thr Arg Lys PTO Arg 665 TTC ATG AGC GCT TGG GC CAA GTG ATC ATA ACC ATT CTG ATT AGT ATG AGC GCT TGG GCC CAA GTG ATC ATA ACC CTC ATG CTA ACA CTA GTG GTG GTG ACC TTG ATC ATT CTG ATT AGT ATG AGC GCT TGG GCC CAA GTG ATC ATA GCC TCC ATT CTG ATT AGT 700 GTA CAG CTA ACA CTA GTG GTG ACC TTG ATC ATA GGA CCC ATG Val Gln Leu Thr Leu Val Val Thr Leu Ile Ile Met Glu Pro Pro Met 720 CCC ATT TTG TCC TAC CCG AGT ATC AGA GAC TAC CTT ATC TTG ATT ACC AGC AAC CTA GTG GTG ACC TTG ATC ATC ATC ATT CTT ACC ATT PTO Ile Leu Ser Tyr Pro Ser Ile Lys Glu Val Tyr Leu Ile Cys Asn 740 ACC AGC AAC CTG GGT GTA GTG GCC CCT GTG GCT TAC AAT GGA CTC CTC TTT TTG TCC TAC CTG AGT ACC TTG ATC ATC ATT AGT TTG CAT TTG TCC AGC GGT GTA GTG GCC CCT GTG GCT TAC CTT ATC TTG AAT ACC AGC AAC CTG GGT GTA GTG GCC CTT GTG GCT TAC CTT ATC TTG ATT 750 ATC ATG AGC TGT ACC TAC TAC TGT GCC CTT AGA ACC CTG AAC GTG CCC CTC TTG ACC AGC AAC CTG GTA GTG GCC CTT AGA ACC CTG AAC GTG CCC GCC TTG ACC AGC AAC CTG GTA GTG GCC CTT AGA ACC CGC AAC GTG CCC GCC TTG ACC ATG ACC TTG ACC TAC TAT GCC TTC ACC ATG GTG CCC GCC TTG ACC AT		Val	Gln	Asp	Glu	Phe 560	Thr	Cys	Arg	Ala	Cys 565	Asp	Leu	Gly	Trp	Trp 570	Pro		
TIGG AGT GAC ATA GAA TCT ATC ATA GCC ATC GCC TTT TCT TGC CTG GGC TTP Ser Asp Ile Glu Ser Ile Ile Ala Ile Ala Phe Ser Cys Leu Gly 599 ATC CTC GTG ACG CTG TTT GTC ACC CTC ATC TTC GTT CTG TAC CGG GAC Ile Leu Val Thr Leu Phe Val Thr Leu Ile Phe Val Leu Tyr Arg Asp 605 ACA CCC GTG GTC AAA TCC TCC AGT AGG GAG CTC TGC TAT ATC ATT CTG Thr Pro Val Val Lys Ser Ser Ser Arg Glu Leu Cys Tyr Ile Ile Leu 625 GCT GGT ATT TTC CTC GGC TAT GTG TGC CCT TTC ACC CTC ATC GCC AAA 2329 Ala Gly Ile Phe Leu Gly Tyr Val Cys Pro Phe Thr Leu Ile Ala Lys 640 CCT ACT ACC ACA TCC TGC TAC CTC CAG CGC CTC CTA GTT GGC CTC TCT Pro Thr Thr Thr Ser Cys Tyr Leu Gln Arg Leu Leu Val Gly Leu Ser 655 TCT GCC ATC TGC TAC TCT GGT TA GTG ACC AAA ACC AAT GGT ACT ACT Ser Ala Met Cys Tyr Ser Ala Leu Val Thr Lys Thr Asn Arg Ile Ala 670 CCC ATC CTG GGT GGC AGC AAG AAG AAG AAG ATC GC CGG AAG CCC AGA Arg Ile Leu Ala Gly Ser Lys Lys Lys Ile Cys Thr Asn Arg Ile Ala 670 TTC ATG AGC GCT TGG GCC CAA GTG ATC ATA GCC CTC ATC TCT ATC GTA CAG CTA ACA CTA GTG GTG ACC TTC ATC ACC CGG AAG CCC AGA Arg Ile Leu Ala Gly Ser Lys Lys Lys Ile Cys Thr Asn Arg Ile Ala 670 TTC ATG AGC GCT TGG GCC CAA GTG ATC ATA GCC TCC ATC TCT ATC TTC ATG AGC GCT TGG GCC CAA GTG ATC ATA GCC TCC ATT CTG ATT AGT Phe Met Ser Ala Trp Ala Gln Val Ile Ile Ala Ser Ile Leu Ile Ser 700 GTA CAG CTA ACA CTA GTG GTG ACC TTG ATC ATC ATC ATC GAG CCT CCC ATG Val Gln Leu Thr Leu Val Val Thr Leu Ile Ile Met Glu Pro Pro Met 720 CCC ATT TTG TCC TAC CCG AGT ATC AAG GAA GTC TAC CTT ATC TGC AAT PPO Ile Leu Ser Tyr Pro Ser Ile Lys Glu Val Tyr Leu Ile Cys Asn 745 ACC AGC AAC CTG GGT GTA GTG GCC CTT GAG GCT TAC ATC GTG CTC TTC ATG AGC CTG ACC TGA GTG GCC CTT GAG GCT TAC ATC GTG CTC TTC ATG AGC TGT ACC TGC GTG ACC TTC ATC CTT ATC TGC ATT TGT TCC TAC CCG AGT ATC AAG GAA GTC TAC CTT ATC TGC ATT 750 ATC ATG AGC TGT ACC TAC TAT GTG GCC CTT GAG ACC CCC AAC GTG CCG CC Thr Ser Asn Leu Gly Val Val Ala Pr Val Gly Tyr Asn Gly Leu Leu 775 ATC ATG AGC TGT ACC TAC TAC TAT GCC TTC AGG ACC CTC AAC GTG CCG GCC Thr Acc AGC ACC TGC		AAC Asn	GCA Ala	GAG Glu	Leu /	ACA Thr	GGC Gly	TGT Cys	GIU	ALO.	ATT Ile	CCT Pro	GTC Val	ur A	-1-	CTT Leu	GAG Glu	•	2137
Trp Ser ASS 11e Glu Ser Ite 11e Ata 11					575	C1.1		እጥሮ	እጥ ል	GCC	ATC	GCC	· TTT	TCT	TGC	CTG	GGC		2185
ILE LEU VAI THE LEU PHE VAI THE LEU ITE PHE VAI COS OF STEED AND THE PHE VAI THE LEU TYPE AND THE PHE LEU GOS OF SEE SEE SEE AND GOS OF STEED AND THE PHE LEU GOS OF SEE SEE AND GOS OF SEE SEE SEE SEE AND GOS OF SEE SEE SEE SEE SEE SEE AND GOS OF SEE SEE SEE SEE SEE SEE SEE SEE SEE SE		Trp	Ser	Asp 590	Ile	Glu	Ser	Ile	595	ALA	TIG	WIG		600	Cyc				
Thr Pro Val Val Lys Ser Ser Ser Arg 610 630 GCT GGT ATT TTC CTC GGC TAT GTG TGC CCT TTC ACC CTC ATC GCC AAA 2329 Ala Gly Ile Phe Leu Gly Tyr Val Cys Pro Phe Thr Leu Ile Ala Lys 640 CCT ACT ACC ACA TCC TGC TAC CTC CAG GCC CTC CTA GTT GGC CTC TCT TTC Thr Thr Ser Cys Tyr Leu Gln Arg Leu Leu Val Gly Leu Ser 665 TCT GCC ATG TGC TAC TCT GCT TTA GTG ACC AAA ACC AAT CGT ATT GCA 2425 Ser Ala Met Cys Tyr Ser Ala Leu Val Thr Lys Thr Asn Arg Ile Ala 670 CGC ATC CTG GGT GGC AGC AAG AAG AAG ATC TGC ACC CGG AAG CCC AGA ATG GTS GS5 TTC ATG AGC GCT TGG GCC CAA GTG ATC ATA GCC TCC ATT CTG ATT AGT 695 TTC ATG AGC GCT TGG GCC CAA GTG ATC ATA GCC TCC ATT CTG ATT AGT 695 TTC ATG AGC GCT TGG GCC CAA GTG ATC ATA GCC TCC ATT CTG ATT AGT 705 TTC ATG AGC GTA ACA CTA GTG GTG ACC TTG ATC ATC ATC ATC ATG GAG CCT CCC ATG 715 GTA CAG CTA ACA CTA GTG GTG ACC TTG ATC ATC ATC ATG GAG CCT CCC ATG 720 CCC ATT TTG TCC TAC CCG AGT ATC AAG GAA GTC TAC CTT ATC TGC AAT 720 CCC ATT TTG TCC TAC CCG AGT ATC AAG GAA GTC TAC CTT ATC TGC AAT 730 CCC ATT TTG TCC TAC CCG AGT ATC AAG GAA GTC TAC CTT ATC TGC AAT 730 CCC ATT TTG TCC TAC CCG AGT ATC AAG GAA GTC TAC CTT ATC TGC AAT 730 CCC ATT TTG TCC TAC CCG AGT ATC AAG GAA GTC TAC CTT ATC TGC AAT 730 CCC ATT TTG TCC TAC CCG AGT ATC AAG GAA GTC TAC CTT ATC TGC AAT 730 ACC AGC AAC CTG GGT GTA GTG GCC CCT GTG GGT TAC AAT GGA CTC CTC TAC TAC TAC TAC TAC TAC TAC TA		ATC Ile	Leu	GTG Val	ACG Thr	CTG Leu	TTT Phe	Val	ACC Thr	CTC Leu	ATC Ile	Lue	AdT.	CTG	TAC Tyr	Arg	Asp		2233
Ala Giy Ile Phe Leu Gly Tyr Val Cys Pro Phe Thr Leu Ile Ala Lys 640 CCT ACT ACC ACA TCC TGC TAC CTC CAG CGC CTC CTA GTT GGC CTC TCT TCT Pro Thr Thr Thr Ser Cys Tyr Leu Gln Arg Leu Leu Val Gty Leu Ser 655 TCT GCC ATG TGC TAC TCT GCT TTA GTG ACC AAA ACC AAT CGT ATT GCA 2425 Ser Ala Met Cys Tyr Ser Ala Leu Val Thr Lys Thr Asn Arg Ile Ala 670 CGC ATC CTG GCT GGC AGC AAG AAG AAG AAC CGG AAG CCC AGA ARG Ile Leu Ala Gly Ser Lys Lys Lys Ile Cys Thr Arg Lys Pro Arg 685 TTC ATG AGC GCT TGG GCC CAA GTG ATC ATA GCC TCC ATT CTG ATT AGT AGT Ile Leu Ala Trp Ala Gln Val Ile Ile Ala Ser Ile Leu Ile Ser 700 CTA CAG CTA ACA CTA CTG GTG ACC TTG ATC ATC ATG GAG CCT CCC ATG PO Met 720 CCC ATT TTG TCC TAC CCG AGT ATC AAG GAA GTC TAC ATG GAG CCT CCC ATG PO Met 725 ACC AGC AAC CTG GGT GTA GTG GCC CCT GTG GGT TAC AAT GGA CTC CTC ATT TTG TCC TAC CTG ATT CTG ATT TTG TCC TAC CTG ATC ATG GAC CTT TAC TGC AAT TAC TTG TAC TTG TAC TTG TAC		Thr	CCC	GTG Val	GTC Val	AAA Lys	Ser	TCC Ser	AGT Ser	AGG	GAG Glu	TRU	TGC Cys	TAT	ATC Ile	110	200	•	2281
TCT GCC ATG TGC TAC TCT GCT TTA GTG ACC AAA ACC AAT CGT ATT GCA Ser Ala Met Cys Tyr Ser Ala Leu Val Thr Lys Thr Asn Arg Ile Ala 675 CGC ATC CTG GCT GGC AGC AAG AAG AAG ATC TGC ACC CGG AAG CCC AGA 2473 Arg Ile Leu Ala Gly Ser Lys Lys Lys Ile Cys Thr Arg Lys Pro Arg 685 TTC ATG AGC GCT TGG GCC CAA GTG ATC ATA GCC TCC ATT CTG ATT AGT 2521 Phe Met Ser Ala Trp Ala Gln Val Ile Ile Ala Ser Ile Leu Ile Ser 710 GTA CAG CTA ACA CTA GTG GTG ACC TTG ATC ATG GAG CCT CCC ATG 715 GTA CAG CTA ACA CTA GTG GTG ACC TTG ATC ATG GAG CCT CCC ATG 725 CCC ATT TTG TCC TAC CCG AGT ATC AAG GAA GTC TAC CTT ATC TGC AAT 740 CCC ATT TTG TCC TAC CCG AGT ATC AAG GAA GTC TAC CTT ATC TGC AAT 740 ACC AGC AAC CTG GGT GTA GTG GCC CCT GTG GGT TAC AAT GGA CTC CTC TAC TAC TAC TAC TAC TAC TAC TA			GGT Gly	ATT Ile	TTC Phe	. Leu	Gly	TAT Tyr	GTG Val	TGC Cys	PLO.	TTC	ACC Thr	CTC Leu	ATC	***	AAA Lys	•	2329
CGC ATC CTG GCT GGC AGC AAG AAG AAG ATC TGC ACC CGG AAG CCC AGA Arg Ile Leu Ala Gly Ser Lys Lys Lys Ile Cys Thr Arg Lys Pro Arg 685 TTC ATG AGC GCT TGG GCC CAA GTG ATC ATA GCC TCC ATT CTG ATT AGT Phe Met Ser Ala Trp Ala Gln Val Ile Ile Ala Ser Ile Leu Ile Ser 700 GTA CAG CTA ACA CTA GTG GTG ACC TTG ATC ATC ATG GAG CCT CCC ATG Val Gln Leu Thr Leu Val Val Thr Leu Ile Ile Met Glu Pro Pro Met 725 CCC ATT TTG TCC TAC CCG AGT ATC AAG GAA GTC TAC CTT ATC TGC AAT Pro Ile Leu Ser Tyr Pro Ser Ile Lys Glu Val Tyr Leu Ile Cys Asn 740 ACC AGC AAC CTG GGT GTA GTG GCC CCT GTG GGT TAC AAT GGA CTC CTC Thr Ser Asn Leu Gly Val Val Ala Pr Val Gly Tyr Asn Gly Leu Leu 755 ATC ATG AGC TGT ACC TAC TAC TAT GCC TTC AAG ACC CGC AAC GTG CCG GCC 11e Met Ser Cys Thr Tyr Tyr Ala Phe Lys Thr Arg Asn Val Pro Ala 765	_	CCT	ACT	ACC	Thr	TCC	TGC	TAC Tyr	CTC	GIR	CGC Arg	CTC Lau	CTA Leu	GTT Val	GT	CTC Leu	TCT	. *	2377
Arg lie Leu Ala Gly Ser Lys Lys Lys lie Cys Hys 116 Cys ATT G95 TTC ATG AGC GCT TGG GCC CAA GTG ATC ATA GCC TCC ATT CTG ATT AGT Phe Met Ser Ala Trp Ala Gln Val lie lie Ala Ser lie Leu lie Ser 710 GTA CAG CTA ACA CTA GTG GTG ACC TTG ATC ATC ATG GAG CCT CCC ATG Val Gln Leu Thr Leu Val Val Thr Leu lie lie Met Glu Pro Pro Met 720 CCC ATT TTG TCC TAC CCG AGT ATC AAG GAA GTC TAC CTT ATC TGC AAT Pro lie Leu Ser Tyr Pro Ser lie Lys Glu Val Tyr Leu lie Cys Asn 745 ACC AGC AAC CTG GGT GTA GTG GCC CCT GTG GGT TAC AAT GGA CTC CTC TAC CTC ATG ALeu Gly Val Val Ala Pr Val Gly Tyr Asn Gly Leu Leu 750 ATC ATG AGC TGT ACC TAC TAT GCC TTC AAG ACC CGC AAC GTG CCG GCC 2713 ATC ATG AGC TGT ACC TAC TAT GCC TTC AAG ACC CGC AAC GTG CCG GCC 2713 ATC ATG AGC TGT ACC TAC TAT GCC TTC AAG ACC CGC AAC GTG CCG GCC 2713	,	TCT Ser	GCC Ala	Met	Cys	TAC Tyr	TCT Ser	GCT Ala	Leu	Val	ACC Thr	AAA Lys	ACC Thr	AU.	CGT	ATT	GCA. Ala	100	2425
Phe Met Ser Ala Trp Ala Gin val lie lie als		CGC	Ile	Leu	GCT Ala	GGC	Ser	Lys	FAR	AAG Lys	ATC Ile	TGC Cys		CGG Arg	AAG Lys	CCC	AGA Arg		2473
GTA CAG CTA ACA CTA GTG GTG ACC TTG ATC ATG GAG CCT CCC ATG Val Gln Leu Thr Leu Val Val Thr Leu Ile Ile Met Glu Pro Pro Met 720 CCC ATT TTG TCC TAC CCG AGT ATC AAG GAA GTC TAC CTT ATC TGC AAT Pro Ile Leu Ser Tyr Pro Ser Ile Lys Glu Val Tyr Leu Ile Cys Asn 745 ACC AGC AAC CTG GGT GTA GTG GCC CCT GTG GGT TAC AAT GGA CTC CTC ASG ASG Leu Gly Val Val Ala Pr Val Gly Tyr Asn Gly Leu Leu 750 ATC ATG AGC TGT ACC TAC TAT GCC TTC AAG ACC CGC AAC GTG CCG GCC 2713 Ile Met Ser Cys Thr Tyr Ala Phe Lys Thr Arg Asn Val Pro Ala 775		Phe	Met	AGC Ser	GCT Ala	TGG Trp	ALA	CAA Gln	GTG Val	ATC	ATA	VT.	TCC	ATT	CTG Leu	ATT	AGT Ser 715		2521
Pro Ile Leu Ser Tyr Pro Ser Ile Lys Glu Val Tyr Det 745 ACC AGC AAC CTG GGT GTA GTG GCC CCT GTG GGT TAC AAT GGA CTC CTC Thr Ser Asn Leu Gly Val Val Ala Pr Val Gly Tyr Asn Gly Leu Leu 750 ATC ATG AGC TGT ACC TAC TAT GCC TTC AAG ACC CGC AAC GTG CCG GCC Ile Met Ser Cys Thr Tyr Ala Phe Lys Thr Arg Asn Val Pro Ala 775 TO THE Leu Ser Tyr Pro Ser Ile Lys Glu Val Tyr Det 745 2665 2713	,			CTA Leu	ACA Thr	· Leu	Aaj	GTG Val	ACC	TTG	TTA	ATC	ATG Met	GAG Glu	CCT Pro			* -	2569
Thr Ser Asn Leu Gly Val Val Ala Pr Val Gly 191 ASN 760 ATC ATG AGC TGT ACC TAC TAT GCC TTC AAG ACC CGC AAC GTG CCG GCC 2713 Ile Met Ser Cys Thr Tyr Ala Phe Lys Thr Arg Asn Val Pro Ala 765 TAG ATC ACC ATG TAC ACC ATG TAC ACC TGC 2761		CCC	ATT	TTG Leu	Ser	Tyr	ccc Pro	AGT Ser	ATC Ile	TAS	GIU	GTC Val	TAC	CTT			AAT Asn		2617
ATC ATG AGC TGT ACC TAC TAT GCC TTC AAG ACC CGC AAC GTG CCG GCC 11e Met Ser Cys Thr Tyr Ala Phe Lys Thr Arg Asn Val Pro Ala 765 770 775 776 777 776 777 776 777 777 778		ACC	AGC Ser	AST	CTG	CCT	GTA Val	GTG Val	TY	. P.E.	r GTG Val	GGT Gly	TAC	AS.		CTC	CTC Leu		2665
2761		ATC	Met	AGC Ser	:	ACC Thi	TAC Tyr	IVI	ATS	TTC Phe	C AAG a Lys	ACC Thr	7.2		GTG Val	CCC.	GCC Ala		2713
		AAG			GAG	GC!	r aas	mac	ን አጥረ	GCG Ala	C TTC	ACC	: ATG : Met	TAC Tyr	ACT Thi	ACC Thi	TGC Cys		2761

				•					:									
	780	, in the second		¥		785					790				, .	795	•	
	` amo	<u>እጥ</u> ር	TICC	СТС	GCT	TTC	GTT	CCC	ATT	TAC	TTT	GGG	AGC	AAC	TAC	AAG		2809
	Ile	Ile	Tro	Leu	Ala	Phe	Val	Pro	Ile	Tyr	Phe	Gly	Ser	Asn	Tyr	Lys		
					800				,	805		. •			810		٠.	•
			·					CMC	1300	 	ACT	GMG	A CG	GTG	GCC	- CTC	Θ	2857
	ATC	ATC	ACT	ACC	TGC	Phe	Ala	Val	Ser	Leu	Ser	Val	Thr	Val	Ala	CTG Leu	. •	2057
	TIE	TIE	1111	815			,,,,,		820			,		825				
								·	-1.						0.00			
	GGG	TGC	ATG	TTT	ACT	CCG	AAG	Mot	TAC	ATC	ATC	ATT	Ala	AAA T.ve	Pro	GAG Glu		2905
	GTĀ	cys	830		THE	PLO	Lys	835	-77-	110	110	- 110	840	273		376		
			•		:			-								•	. ** .	٠.
	AGG	AAC	GTC	CGC	AGT	GCC	TTC	ACG	ACC	TCT	GAT	GTT	GTC	CGC	ATG	CAC	•	2953
	_		Val	Arg	Ser	Ala	Phe 850	Thr	Thr	Ser	Asp	855	Val	Arg	Met	HIS		
	•• 1	845					. 050					7				. •	÷	
	GIC	GGT	GAT	GGC	AAA	CTG	CCG	TGC	CGC	TCC	AAC	ACC	TTC	CTC	AAC	ATT		3001
	Val	Gly	Asp	GLY	Lys	Leu	Pro	Cys	Arg	Ser	Asn	Thr	Phe	Leu	Asn	Ile		-
	860	-			,	865				•	870	. •		•		875	٠.,	Θ.
	יייי	CGG	AGA	AAG	AAG	ccc	GGG	GCA	GGG	AAT	GCC	AAT	TCT	AAC	GGC	AAG		3049
	Phe	Arg	Arg	Lys	Lys	Pro	Gly	Ala	Gly	Asn	Ala	Asn	Ser	. Asn	Gly	Lys		
			_	· ·	880		• .	*		885		*	. :		890			
	mem	CIDO	me a	TCC	- The state of the	CAA	CCA	GGT	GGA	AGA	CAG	GCG	CCC	AAG	GGA	CAG		3097
	Ser	Val	Ser	Tro	Ser	Glu	Pro	Gly	Gly	Arg	Gln	Ala	Pro	Lys	Gly	Gln	•	
				895	0			_	900	_			*	905				
						~~~	MAR	cmc	C2 C	CTIC	226	A CC	330	GAG	ACG	GCC		3145
	CAC	GIG	TGG	CAG	Arm	Leu	Ser	Val	His	Val	Lvs	Thr	AAC	Glu	Thr	Ala		
	uis	val	910	91				915					920					
	-				· 						. ,		3.00	msá	G3.3		**	2102
	TGT	AAC	CAA	ACA	GCC	GTA	ATC	AAA	Dro	Len	Thr	TVS	AGT Ser	TVY	Gln	GIV		3193
	cys	<b>ASN</b> 925	GTU.	THE	VIG	VAI	930	בענם				935		-1-		1		
	:			-									• •	• • •	*	· .	· · · · · · · · · · · · · · · · · · · ·	
	TCT	GGC	AAG	AGC	CTG	ACC	TTT	TCA	GAT	GCC	AGC	ACC	AAG	ACC	CTT	TAC		3241
			Lys	Ser	Leu	<b>TDF</b> 945	Pne	Ser	ASP	ATA	950	THE	Lys	THE	rėu	955	٠٠.	10
	940			-	•				* `.	· .		•••	•	•		•	•	
	AAT	GTG	GAA	GAA	GAG	GAC	TAA	ACC	CCT	TCT	GCT	CAC	TTC	AGC	CCT	CCC	. ;	3289
	Asn	Val	Glu	Glu	Glu	Asp	Asn	Thr	Pro	Ser	Ala.	His	Phe	Ser	Pro. 970	Pro		
	*				960					965				•	370	1 .	7	٠
	AGC	AGC	CCT	TCT	ATG	GTG	GTG	CAC	CGA	CGC	GGG	CCA	CCC	GTG	GCC	ACC		3337
	Ser	Ser	Pro	Ser	Met	Val	Val	His	Arg	Arg.	Gly	Pro	Pro	Val	Ala	Thr		
	•		•	975		. •			980		<b>.</b>	-	•	985		,		_
	141	003	- COM	CALC	CCA	CCC	САТ	CTG	ACC	GCA	GAA	GAG	ACC	CCC	CTG	TTC		3385
•	ACA Thr	Pro	Pro	Lau	Pr	Pro	His	Leu	Thr	Ala	Glu	Glu	Thr	Pro	Leu	Ph		<del>-</del>
			990		- <del>-</del>			995		• •			1000	)·				٠.
	•											COM	COM	~; ~;	000	CNC	* • • <b>•</b>	7/77
	CTG	GCT	GAT	TCC	GIC	ATC	Pro	AAG Tue	<b>ਚ</b> ਫ਼ਿ <b>ਹ</b> ਹੀ ਦ	TIG	Pr	Pr	CCT Pr	Lau	Pro	Gln		3433
		A18 1005		Jer	AGT	TT2	1010		<b>1</b>			1015					•	
			•					,					•	•		-	• •	<u>/</u>

	٠.					1 1						·				* *	
G	ln (	Gln	CCA Pro	CAG Gln	CAG Gln	CCG Pro	PLO	CCT	CAG Gln	CAG Gln	CCC Pro 1030		CAG Gln	CAG Gln	Pro	AAG Lys 1035	3481
	020 CC er		ATG Met	GAC Asp	GIN	CTG Leu		GGC Gly	GTA Val	GTC Val		AAC Asn	TTC Phe	GGT Gly	TCG Ser 1050	GGG Gly	3529
				• 8 .	1040	•		, •			•		3.03	CCA	GGA	AAC	3577
A	TT	CCA Pro	GAT Asp	TTC Phe 105	His	GCG Ala	GTG Val	CTG Leu	GCA Ala 106	GIJ	CCG	Gly	Thr	Pro 106	Gly	Asn	
2	GC	CTG	CGC	TCT	CTG	TAC	CCG	CCC	CCG	CCT	CCG	CCG	CAA	CAC	CTG	CAG Gln	3625
5	er	Leu	Arg	Ser	Leu	Tyr	Pro	Pro 107	PLO	Pro	Pro	Pro	108	0	Dec	• .	∌
	\TG {et	CTG J. au	CCC	CTG Leu	CAC His	CTG Leu	Ser	THE	TTC Phe	CAG Gln	GAG Glu	GAG Glu 109		ATC Ile	TCC	CCT Pro	3673
		108	5				TOS	U					•				3721
1	CT Pro	Gly	GAG Glu	GAC Asp	ATC Ile	GAT Asp 110	ASP	GAC Asp	AGT Ser	GAG Glu	AGA Arg 111		Lys	Leu	Leu	CAG Gln 1115	3/21
		•	GTG Val	TAC	GIA	Arg	GAA	GGG Gly	AAC Asi			GAA Glu	GAT Asp	GAA Glu	TTG Leu 113	GAA Glu O	3769
					112	U	•							•			3817
	GAG Glu	GAG Glu	GAC Glu	GAC Asp 113	Lev	CCC Pro	ACI Thi	A GCC	C AGO A Sei 114	. Dy .	CTG Leu	Thi	Pro	GAG Glu 114	Asp 5	TCT Ser	
a .	CCT Pro	GC0 Ala	Lev	ACG Thi		CC:	r TC	E PL	r TT o Pho	c cgi	A GAT J Asp	TCC Sei	C GTC Val	GCC L Ala	TCI Ser	GGC Gly	3865
	AGC Ser	Se	r Va	·-	AG Se	r TC r Se	C CC r Pr	C GT		T GA	G TCC	G GT(		C TGC	C ACC	C CCT	3913
	Pro	As		A ACC	TA TY	L AT			C AT 1 Il	T CT e Le	G AGG	,	C TA	C AA	G CA	A AGC n Ser 1195	3961
	118							CMC	יייי	ምርር	GGGC		GG A	GTGC	GCAT	G ; ; ;	4013
٠	TCI	TC Se	C AC r Th	r Lei	G TA	GTGI	GIGI	GIG	1010	100			8	• ,		•	-00
•							CNOT	יביירר ז	.ac (	كست	CAGA	A AT	GTGT	AGAA	AGC	AGGGTGA	4073
	GAC	GAAG	CCAG	AGA	TGCC	AAG .	CAGI	GT CH			***	2 22	አል አጥ	· GCTG	CGG	CTGCCTT	4133
	GG	BATG	GGGA	TGG	AGGA	CCA	CGGI	CTGC	AG G	GAAG		w	CACC	المالة	TGA	CTGCCTT CTCTTGA	4193
	AA	AGAA	GGAG	AGG	GACG	ATG	CCAP	CTG	AAC A	\GTG(	TCCT	ان ن	, LAGG	17 W 2 2	mm-	CTCTTGA	4253
															110	CATATGG	4300
	TA'	TGTA	ACTI	TTA	TCGA	AAA	AAAI	AAAA	AAA 2	LAAA	AAAA	A A	LAAA	<b>LA</b>			4300

## (2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1199 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
- Met Val Arg Leu Leu Ile Phe Phe Pro Met Ile Phe Leu Glu Met
- Ser Ile Leu Pro Arg Met Pro Asp Arg Lys Val Leu Leu Ala Gly Ala
- Ser Ser Gln Arg Ser Val Ala Arg Met Asp Gly Asp Val Ile Ile Gly
- Ala Leu Phe Ser Val His His Gln Pro Pro Ala Glu Lys Val Pro Glu
  50 60
- Arg Lys Cys Gly Glu Ile Arg Glu Gln Tyr Gly Ile Gln Arg Val Glu 65
- Ala Met Phe His Thr Leu Asp Lys Ile Asn Ala Asp Pro Val Leu Leu 90 95
- Pro Asn Ile Thr Leu Gly Ser Glu Ile Arg Asp Ser Cys Trp His Ser 100
- Ser Val Ala Leu Glu Gln Ser Ile Glu Phe Ile Arg Asp Ser Leu Ile 115
- Ser Ile Arg Asp Glu Lys Asp Gly Leu Asn Arg Cys Leu Pro Asp Gly
- Gln Thr Leu Pro Pro Gly Arg Thr Lys Lys Pro Ile Ala Gly Val Ile 145 2. 150
- Gly Pro Gly Ser Ser Val Ala Ile Gln Val Gln Asn Leu Leu Gln 175
- Leu Phe Asp Ile Pro Gln Ile Ala Tyr Ser Ala Thr Ser Ile Asp Leu 185
- Ser Asp Lys Thr Leu Tyr Lys Tyr Phe Leu Arg Val Val Pro Ser Asp 205
- Thr Leu Gln Ala Arg Ala Met Leu Asp Ile Val Lys Arg Tyr Asn Trp
  210 215
- Thr Tyr Val S r Ala Val His Thr Glu Gly Asn Tyr Gly Glu Ser Gly 240

	0		- 9							,,	, , , , , , , , , , , , , , , , , , ,					
	Mez	Asp	Ala	Phe	Lys 245	Glu	Leu	Ala	Ala	Gln 250	Glu	Gly	Leu	Cys	Ile 255	Ala
	His	Ser	Asp	Lys 260	Ile	Tyr	Ser	Asn	Ala 265	Gly	Glu	Lys	Ser	Phe 270	Asp	Arg
	Leu	Leu	Arg 275	Lys	Leu	Arg	Glu	Arg 280	Leu	Pro	Lys	Ala	Arg 285	Val	Val	Val
	<del></del>	290	• ,	1	Gly		295					300				. •
	305		1	١.	•	310	*				313					
					Glu 3 <b>25</b>		•	•								
	٠.			340				ï	345							
	•	,	355		Leu	٠.	. •	200	•	*				1		
	• .	370	).		Gln		375		•			300		•		
•	385		. • •		) Asn	390		•								
				, v .	Val 405		•	•	•	410	,	÷		·		
				420	)				425					:		Cys
		:	.43	5				440								Arg
•		45	0	,	_•.		495								-	Gly
	465	5	•	•		470	, Y	1			, 47.	•		'		480
		:	-	, ,	485	<b>)</b> .				771						His
				50	<b>o</b> ,				. 50.	-	•	•		*	/-	ile
	•	•	51	.5				520	9					_	•	Cys
	Le	u Ly 53		y Gl	n Ile	e Lys	53	1 Il 5	e ·Ar	g Ly	s Gl	y Gli 54	u Vai	1 S	r cy	s Cys

					•	٠.						_			-1.	
545		,	. 0	Ala	つつひ										•	
Thr	Cys	Arg	Ala	Cys 565	Asp	Leu	Gly	Trp	Trp 570	Pro	Asn	Ala	Glu	Leu 575	Thr	
Gly	Cys	Glu	Pro 580	Ile	Pro	Val	Arg	Tyr 585	Leu	Glu	Trp	Ser	<b>Asp</b> 590	Ile	Glu	
Ser	Ile	Ile 595	Ala	Ile.	Ala	Phe	Ser 600	Суз	Leu	Gly	Ile	Leu 605	Val	Thr	Leu	٠.
Phe	Val	Thr	Leu	Ile	Phe	Val 615	Leu	Tyr	Arg	Asp	Thr 620	Pro	Val	Val	Lys	
Ser	Ser	Ser	Arg	Glu	Leu 630	Cys	Tyr	Ile	Ile	Leu 635	Ala	Gly	Ile	Phe	Leu 640	
625 Gly	Tyr	val	Сув	Pro 645	-	Thr	Leu	Ile	Ala 650	Lys	Pro	Thr	Thr	Thr 655	Ser	
Cys	Tyr	Leu	Gln	Arg	Leu	Leu	Val	Gly 665	Leu	Ser	Ser	Ala	Met 670	Cys	Tyr	
Set	Ala	Lev	660 Val		Lys	Thr	Asn 680	20		Ala	Arg	Ile 685	Leu	Ala	Gly	
Sei	. Lys	675 Lys		Ile	Cys	Thr 695	Arg	Lys	Pro	Arg	Phe 700	Met	Ser	Ala	Trp	
	690	)	1		Ala	Ser					Val		•		Leu 720	
705	5				./10				Pro	Met					Tyr	
				7.25	١.										Gly	
	٠,		. 740	)					*	•					Thr	,
		75	5	<b>2</b> -			, • •	,		•	,			1	1 Ala	
	. 77	0				,,,		•		. `			•			
· 78	5	. 1		٠.	126	,									Ala 800	
		- 5		80:	<del>)</del>			•		_	-		•		r Cys 5	
			82	U			. •		•				4	_	e Thr	
Pr	o Ly	s Me 83	t Ty	r Il	e Il	e Il	e Ala 84	a Ly	s Pr	o Gl	u Ar	g As: 84	n Va 5	1 Ar 	g Ser	•

Ala	Phe 850	Thr	Thr	Ser	Asp	Val 855	Val	Arg	Met	His	Val 860	Gly	Asp	Gly	Lys
Leu 865	Pro	Cys	Arg	ser	Asn 870	Thr	Phe	Leu	Asn	Ile 875	Phe	Arg	Arg	Lys	Lys 880
Pro	Gly	Ala	Gly	Asn 885	Ala	Asn	Ser	Asn	Gly 890	Lys	Ser	Val	Ser	Trp 895	Ser
Glu	Pro	Gly	Gly 900	Arg	Gln	Ala	Pro	Lys 905	Gly	Gln	His	Val	Trp 910	Gln	Arg
Leu	Ser	Val 915	His	Val	Lys	Thr	Asn 920	Glu	Thr	Ala	Cys	Asn 925	Gln	Thr	Ala
Val	Ile 930	Lys	Pro	Leu	Thr	Lys 935	Ser	Tyr	Gln	Gly	Ser 940	Gly	Lys	Ser	Leu
Thr. 945	Phe	Ser	Asp	Ala	Ser 950	Thr	Lys	Thr	Leu	Tyr 955	Asn	Val	Glu	Glu	Glu 960
Asp	Asn	Thr	Pro	Ser 965	Ala	His	Phe	Ser	Pro 970	Pro	Ser	Ser	Pro	Ser 975	Met
Val	Val	His	Arg 980	Arg	Gly	Pro	Pro	Val 985	Ala	Thr	Thr	Pro	Pro 990	Leu	Pro
Pro	His	Leu 995		Ala	Glu	Glu	Thr 100	Pro 0	Leu	Phe	Leu	100	Asp 5	Ser	Val
Ile	Pro 101		Gly	Leu	Pro	Pro 101	Pro 5	Leu	Pro	Gln	Gln 102	Gln O	Pro	Gln	Gln
Pro	Pro	Pro	Gln	Gln	Pro 103	Pro 0	Gln	Gln	Pro	Lys 103	Ser 5	Leu	Met	. Asp	Gln 1040
Lev	Glr	Gly	val	Val 104	Thr 5	Asn	Phe	Gly	Ser 105	Gly 50	Ile	Pro	) Asp	) Phe 105	His 5
Ala	Val	Let	1 Ala 105	Gly 60	Pro	Gly	Thr	Pro 106	Gly 55	) Asr	s Ser	Lev	107	g Ser 10	Leu
Туз	Pro	) Pro	Pro 75	Pro	) Pro	Pro	Glr 108	His BO	Let	i Glr	ı Met	108	Pro	Leu	His
Let	1 Set	Th:	r Phe	e Glr	ı Glu	Gl:	ı Ser 95	: Ile	e Se	r Pro	) Pro	Gly	, Gli	ı Asp	Ile
	o Asj 05	o <b>As</b> j	Sea	c Glu	a Arg	Phe .0	a Lys	i Leu	ı Le	u Gl:	n Glu 15	ı Pho	e Va.	l Tyi	Glu 1120
Arc	g Gl	ı Gl	y Asi	112	r Glu 25	Gl	ı Ası	Gl:	1 Le 11	u Gli 30	u Glu	ı Glı	u Gl	u Ası 11:	Leu 35
Pr	Th	r Al	a Se:	r Lys 40	s Lev			114	43			o Ala	11	u Thi 50	r Pr
••				•		SU	BSTI	TUT	e St	<b>EET</b>			-		

78 Pro Ser Pro Phe Arg Asp Ser Val Ala Ser Gly Ser Ser Val Pro Ser 1160 ູ 1155 Ser Pro Val Ser Glu Ser Val Leu Cys Thr Pro Pro Asn Val Thr Tyr 1180 1175 1170 Ala Ser Val Ile Leu Arg Asp Tyr Lys Gln Ser Ser Ser Thr Leu 1190 1185 (2) INFORMATION FOR SEQ ID NO:3: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 35 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 'ii) MOLECULE TYPE: cDNA (vii) IMMEDIATE SOURCE: (B) CLONE: ZC775 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3: GCTAGCATAA CCCCTTGGGG CCTCTAAACG GGTCT (2) INFORMATION FOR SEQ ID NO:4: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 43 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: CDNA (vii) IMMEDIATE SOURCE: (B) CLONE: ZC776

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CTCAAGACCC GTTTAGAGGC CCCAAGGGGT TATGCTAGCT GCA

- (2) INFORMATION FOR SEQ ID NO:5:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 40 base pairs
    - (B) TYPE: nucleic acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: CDNA

(vii)	IMMED	IATE	SC	URCE:
•				ZC777

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:
TGAGGGGTTT TTTGCTGAAA GGAGGAACTA TGCGGCCGCA

40

## (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 40 base pairs
  - (B) TYPE: nucleic acid.
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (Vii) IMMEDIATE SOURCE: (B) CLONE: ZC778

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6: AGCTTGCGGC CGCATAGTTC CTCCTTTCAG CAAAAAACCC

40

## (2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: CDNA
- (vii) IMMEDIATE SOURCE: (B) CLONE: ZC1751
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AATTCTGTGC TCTGTCAAG

19

## (2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucl ic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: CDNA

100

CCCGGGCTCC CGGCAGTGCG AGCAGCTAAG GGCTGGCCGC CGCCTCCCTG AGCTCCCCG GMGCAGCCGA CCCCTGGTCG CGGCGTTCAC CTCGCCGATG CGCGGTTGGT AGGAGTGĀCC GGAGCCATTC TCTCCTCGTT GATAAGATTC CCTACCAGGA TAGGAGCCTA TCTCCCTTTY CACAGCAGGA CACAGAAATC TGGCCTTCAG TACTTTGGGA AAAGGATCTG AGACCTCCTG GAGCTCTGAC CACTGGCTGT CATCTGTGGC TCTGGCCTGT GTGGGCCACT GAGCTCTĀCT 360 CAAACATTAA AGAGGAGGAG GGGAGATCTG TGGAATGGGC CACCCCGTTG GCCTGCTGCA TTACTGAACC TGCGCTGTCC ACACGTGCCC AGATCATGGG ACCCAGGGCC TGCTAGGGCT AGGAGCGGGG CCCAGTATTC ATGGGTCTCT AGGCCTTTCC GAA ATG TCC GGG AAG Met Ser Gly Lys GGA GGC TGG GGC TGG TGG GCC CGG CTG CCC CTC TGC CTA CTC CTC Gly Gly Trp Ala Trp Trp Ala Arg Leu Pro Leu Cys Leu Leu Leu AGC CTT TAT GCC CCC TGG GTG CCT TCA TCC TTG GGA AAG CCC AAG GGT Ser Leu Tyr Ala Pro Trp Val Pro Ser Ser Leu Gly Lys Pro Lys Gly CAC CCC CAC ATG AAC TCT ATC CGA ATT GAC GGG GAC ATC ACA CTG GGA His Pro His Met Asn Ser Ile Arg Ile Asp Gly Asp Ile Thr Leu Gly GGC CTG TTT CCC GTC CAC GGC CGT GGC TCT GAG GGT AAG GCC TGC GGG Gly Leu Phe Pro Val His Gly Arg Gly Ser Glu Gly Lys Ala Cys Gly GAG CTG AAG AAG GAG AAA GGC ATC CAC CGC CTG GAG GCC ATG CTG TTT Glu Leu Lys Lys Glu Lys Gly Ile His Arg Leu Glu Ala Met Leu Phe GCC CTG GAC CGC ATC AAC AAT GAC CCG GAC CTA CTG CCC AAC ATC ACG Ala Leu Asp Arg Ile Asn Asn Asp Pro Asp Leu Leu Pro Asn Ile Thr

FIG. 8A.

ITG GGC GCC CGC ATT CTG GAC ACC IGC ICG AGG GAC ACC CAC GCC CTG Leu Gly Ala Arg Ile Leu Asp Thr Cys Ser Arg Asp Thr His Ala Leu GAG CAG TCA CTG ACC TTT GTG CGG GCG CTC ATC GAG AAG GAC GGC ACG Glu Gln Ser Leu Thr Phe Val Arg Ala Leu Ile Glu Lys Asp Gly Thr GAG GTC CGC TGC GGC AGG CGG GGC CCC ATC ATC ACC AAG CCC GAA Glu Val Arg Cys Gly Arg Arg Gly Pro Pro Ile Ile Thr Lys Pro Glu CGA GTG GTG GTC ATT GGA GCT TCG GGG AGC TCC GTC TCG ATC ATG Arg Val Val Gly Val Ile Gly Ala Ser Gly Ser Ser Val Ser Ile Met 150 155 GTG GCC AAC ATC CTC CGC CTC TTC AAG ATC CCT CAG ATC AGC TAT GCC Val Ala Asn Ile Leu Arg Leu Phe Lys Ile Pro Gln Ile Ser Tyr Ala TCC ACG GCC CCT GAC TTG AGT GAC AGC CGC TAT GAC TTC TTC TCC Ser Thr Ala Pro Asp Leu Ser Asp Asn Ser Arg Tyr Asp Phe Phe Ser CGG GTG GTG CCC TCA GAC ACA TAC CAG GCC CAG GCC ATG GTG GAT ATT Arg Val Val Pro Ser Asp Thr Tyr Gln Ala Gln Ala Met Val Asp Ile 205 GTC CGA GCC CTC AAG TGG AAC TAT GTG TCC ACA CTG GCC TCA GAG GGC Val Arg Ala Leu Lys Trp Asn Tyr Val Ser Thr Leu Ala Ser Glu Gly AGC TAC GGT GAG AGT GGT GTG GAG GCC TTT ATC CAG AAG TCC CGA GAG Ser Tyr Gly Glu Ser Gly Val Glu Ala Phe Ile Gln Lys Ser Arg Glu AAC GGA GGT GTG TGC ATT GCC CAG TCG GTG AAG ATT CCA CGG GAA CCC Asn Gly Gly Val Cys Ile Ala Gln Ser Val Lys Ile Pro Arg Glu Pro 250 245 🐇 260 AAG ACG GGG GAG TTC GAC AAG ATC ATC AAA CGC CTA CTG GAA ACA TCC Lys Thr Gly Glu Phe Asp Lys Ile Ile Lys Arg Leu Leu Glu Thr Ser 265 270 275

FIG. 8B.

25/32 AAT GCC AGG GGT ATC ATC TTT GCC AAC GAG GAT GAC ATC AGG AGG Asn Ala Arg Gly Ile Ile Ile Phe Ala Asn Glu Asp Asp Ile Arg Arg 280 285 GTG TTG GAG GCA GCT CGC AGG GCC AAC CAG ACC GGC CAC TTC TTT TGG Val Leu Glu Ala Ala Arg Arg Ala Asn Gln Thr Gly His Phe Phe Trp 300 ATG GGT TCT GAT AGC TGG GGC TCC AAG AGT GCC CCT GTG CTG CGC CTT Met Gly Ser Asp Ser Trp Gly Ser Lys Ser Ala Pro Val Leu Arg Leu 310 GAG GAG GTG GCC GAG GGC GCA GTC ACC ATT CTC CCC AAG AGG ATG TCT Glu Glu Val Ala Glu Gly Ala Val Thr Ile Leu Pro Lys Arg Met Ser GTT CGA GGG TTC GAC CGA TAC TTC TCC AGC CGC ACG CTG GAC AAC AAC Val Arg Gly Phe Asp Arg Tyr Phe Ser Ser Arg Thr Leu Asp Asn Asn AGG CGC AAC ATC TGG TTT GCC GAG TTC TGG GAG GAC AAC TTC CAT TGC Arg Arg Asn Ile Trp Phe Ala Glu Phe Trp Glu Asp Asn Phe His Cys **360**, 365 3/0 AAG TTG AGC CGC CAC GCG CTC AAG AAG GGA AGC CAC ATC AAG AAG TGC Lys Leu Ser Arg His Ala Leu Lys Lys Gly Ser His Ile Lys Lys Cys 375 ACC AAC CGA GAG CGC ATC GGG CAG GAC TCG GCC TAT GAG CAG GAG GGG Thr Asn Arg Glu Arg Ile Gly Gln Asp Ser Ala Tyr Glu Gln Glu Gly 395 390 AAG GTG CAG TTC GTG ATT GAC GCT GTG TAC GCC ATG GGC CAC GCG CTG Lys Val Gln Phe Val Ile Asp Ala Val Tyr Ala Met Gly His Ala Leu 405 410 415 CAC GCC ATG CAC CGT GAC CTG TGT CCC GGC CGC GTA GGA CTC TGC CCT His Ala Met His Arg Asp Leu Cys Pro Gly Arg Val Gly Leu Cys Pro 435 CGC ATG GAC CCC GTG GAT GGC ACC CAG CTG CTT AAG TAC ATC AGG AAC Arg Met Asp Pro Val Asp Gly Thr Gln Leu Leu Lys Tyr Ile Arg Asn 445 450 440

FIG. 8C.

26/32 GTC AAC TTC TCA GGC ATT GCG GGG AAC CCT GTA ACC TTC AAT GAG AAC Val Asn Phe Ser Gly Ile Ala Gly Asn Pro Val Thr Phe Asn Glu Asn 460 GGA GAC GCA CCG GGG CGC TAC GAC ATC TAC CAG TAC CAA CTG CGC AAT Gly Asp Ala Pro Gly Arg Tyr Asp Ile Tyr Gln Tyr Gln Leu Arg Asn GGC TCG GCC GAG TAC AAG GTC ATC GGC TCG TGG ACA GAC CAC CTG CAC Gly Ser Ala Glu Tyr Lys Val Ile Gly Ser Trp Thr Asp His Leu His CTC AGA ATA GAG CGG ATG CAG TGG CCA GGG AGT GGC CAG CAG CTG CCG Leu Arg Ile Glu Arg Met Gln Trp Pro Gly Ser Gly Gln Gln Leu Pro 505 CGC TCC ATC TGC AGT CTG CCC TGC CAG CCC GGG GAG CGA AAG AAG ACT Arg Ser Ile Cys Ser Leu Pro Cys Gln Pro Gly Glu Arg Lys Lys Thr 520 GTG AAG GGC ATG GCT TGC TGC TGG CAC TGC GAG CCC TGC ACC GGG TAC Val Lys Gly Met Ala Cys Cys Irp His Cys Glu Pro Cys Thr Gly Tyr 540 CAG TAC CAA GTG GAC CGC TAC ACC TGT AAG ACC TGC CCC TAC GAC ATG Gln Tyr Gln Val Asp Arg Tyr Thr Cys Lys Thr Cys Pro Tyr Asp Met 555 CGG CCC ACA GAG AAC CGC ACG AGC TGC CAG CCC ATC CCC ATC GTC AAG Arg Pro Thr Glu Asn Arg Thr Ser Cys Gln Pro Ile Pro Ile Val Lys 570 TTG GAG TGG GAC TCG CCG TGG GCC GTG CTG CCC CTC TTC CTG GCC GTG Leu Glu Trp Asp Ser Pro Trp Ala Val Leu Pro Leu Phe Leu Ala Val GTG GGC ATC GCC GCC ACG CTG TTC GTG GTG GTC ACG TTT GTG CGC TAC Val Gly Ile Ala Ala Thr Leu Phe Val Val Val Thr Phe Val Arg Tyr 605 600 AAC GAT ACC CCC ATC GTC AAG GCC TCG GGC CGG GAG CTG AGC TAC GTG Asn Asp Thr Pro Ile Val Lys Ala Ser Gly Arg Glu Leu Ser Tyr Val 620 615

FIG. 8D.

,	•					,									٠.	
: ,	CTG Leu				ATC Ile				TAC						ATG	
	GCA Ala 645	GAG G1 u	CCG Pro	GAC Asp	CTG Leu	GGG Gly 650	ACC Thr	TGT Cys	TCG Ser	CTC Leu	CGC Arg 655	CGC Arg	ATC Ile	TTC Phe	CTA	2443 GGG Gly 660
•	CTC Leu	GGC Gly	ATG Met	AGC Ser	ATC Ile 665	AGC Ser	TAC Tyr	GCG Ala	GCC Ala	CTG Leu 670	CTG Leu	ACC Thr	AAG Lys	ACC Thr	AAC	2491 CGC Arg
	ATT Ile	TAC Tyr	CGC Arg	ATC Ile 680	TTT Phe	GAG G1u	CAG Gln	GGC Gly	AAA Lys 685	C <b>GG</b> Arg	TCG Ser	GTC Val	AGT Ser	GCC Ala 690	CCG	2539 CGT Arg
-7	TTC Phe	ATC Ile	AGC Ser 695	CCG Pro	GCC Ala	TCG Ser	CAG Gln	CTG Leu 700	GCC Ala	ATC Ile	ACC Thr	TTC Phe	ATC Ile 705	CTC Leu	ATC	2587 TCC Ser
	Leu	CAG Gln 710	CTG Leu	CTC Leu	GGC Gly	ATC Ile	TGC Cys 715	GTG Val	TGG Trp	TTC Phe	GTG Val		GAC Asp		TCC	
		Val.			TTC Phe										TTT Phe	2683 GCC Ala 740
, "					AAG Lys 745										TGC Cys 755	
	_	GGC Gly	TAC	AGC Ser 760	ATG Met	CTG Leu	CTG Leu	ATG Met	GTC Val 765	ACG Thr	TGT Cys	ACT Thr	GTG Val			
		ACC Thr	- !	GGC Gly	GTG Val	CCC Pro	GAG Glu	ACC Thr 780	TTC Phe	AAC Asn	GAG G1u	GCC Ala	AAG Lys 785	CCC Pro	ATC Ile	Gly
	TTC Phe	ACC Thr 790	ATG Met	TAC Tyr	ACC Thr	ACC Thr	TGC Cys 795	ATT	GTC Val	TGG Trp	CTG Leu	GCC Ala 800	TTC Phe	ATC Ile	CCC	2875 ATC Ile

FIG. 8E.

TIT TIT GGC ACC TCA CAG TCA GCC GAC AAG CTG TAC ATC CAG ACA ACC Phe Phe Gly Thr Ser Gln Ser Ala Asp Lys Leu Tyr Ile Gln Thr Thr 805

ACA CTG ACG GTC TCC GTG AGT CTG AGC GCT TCA GTG TCC CTG GGG ATG Thr Leu Thr Val Ser Val Ser Leu Ser Ala Ser Val Ser Leu Gly Met 825

CTC TAC ATG CCC AAA GTC TAC ATC ATC CTC TTC CAC CCG GAG CAG AAC Leu Tyr Met Pro Lys Val Tyr Ile Ile Leu Phe His Pro Glu Gln Asn 840

GTG CCC AAG CGC AAG CGC AGT CTC AAA GCC GTG GTC ACC GCC ACC Val Pro Lys Arg Lys Arg Ser Leu Lys Ala Val Val Thr Ala Ala Thr 855

ATG TCC AAC AAG TTC ACA CAG AAG GGC AAC TTC AGG CCC AAT GGG GAA Met Ser Asn Lys Phe Thr Gln Lys Gly Asn Phe Arg Pro Asn Gly Glu 870

GCC AAA TCA GAG CTG TGT GAG AAC CTG GAG ACC CCA GCG CTG GCT ACC Ala Lys Ser Glu Leu Cys Glu Asn Leu Glu Thr Pro Ala Leu Ala Thr 885

AAA CAG ACC TAC GTC ACC TAC ACC AAC CAT GCC ATC TAGCCGGGCC Lys Gln Thr Tyr Val Thr Tyr Thr Asn His Ala Ile 905

GCGGAGCCAA GCAGGCTAAG GAGCCACAAC CTCTGAGGAT GGCACATTGG GCCAGGGCCG
TTCCCGAGGG CCCTGCCGAT GTCTGCCCGC CTCCCGGGCA TCCACGAATG TGGCTTGGTG
CTGAGGACAG TAGAGACCCC GGCCATCACT GCTGGGCAAG CCGTGGTGGG CAACCAGAGG
AGGCCGAGTG GCTGGGGCAG TTCCAGGTTA TGCCACACAC AGGTCTTCCT TCTGGACCAC
TGTTGGCCCA GCCCCAAAGC ACAGGGGCTC GGTCTCCAGA GCCCAGCCCT GGCTTCCTCT
CCTTCCTCCT GCCTCCGTCT GTCCTGTGGG TGACCCCGGT TGGTCCCTGC CCCGTCTTTA
CGTTTCTCTT CCGTCTTTGC TCTGCATGTG TTGTCTGTTT GGGCCCTCT CTTCCATATT

FIG. 8F.

TITICCATTCT GCTCCTGGCC TTCCCCTGCC ATCTGCCCTG CCCCCTGCCC CTCCTCCCTG

AGCTGCCCCA TCCCCGCCAT CATTTTCTCT TCTGTTCCCC CTCGATCTCA TTTCCTACCA

GCCTTCCCCC TACTTGGCTT CATCCACCAA CTCTTTCACC ACGTTGCAAA AGAGAAAAAA

AAAGGGGGGG GGGAATCACC CCCTACAAAA AAGCCCAAAC AAAAACTAAT CTTGAGTGTG

TTTCGAAGTG CTGCGTCCTC CTGGTGGCCT GTGTGTCCCT GTGCCTGCAG CCTGTCTGCC

CGCCCTACCC GTCTGCCGTG TGTCCTGCCC CCCCGCCTG CCCGCCTTGC CCTTCCTGCT

AACGACACGG AGTTCAGTGC CTGGGTGTTT GGTGATGGTC TCTGATGTGT AGCATGTCTG

TTTTTATACC GAGAACATTT CTAATAAAGA TAAACACATG GTTTTGC

FIG. 8G.

CCCAACATCA CGTTGGGCGC CCGCATTCTG GACACCTGCT CGAGGGACAC CCACGCCCTG GAGCAGTCAC TGACCTTTGT GCGGGCGCTC ATCGAGAAGG ACGGCACGGA GGTCCGCTGC 180 GGCAGGCGGG GCCCGCCCAT CATCACCAAG CCCGAACGAG TGGTGGGTGT CATTGGAĞČŤ 240 TCGGGGAGCT CCGTCTCGAT CATGGTGGCC AACATCCTCC GCCTCTTCAA GATCCCTCAG ATCAGCTATG CCTCCACGGC CCCTGACTTG AGTGACAACA GCCGCTATGA CTTCTTCŤČČ CGGGTGGTGC CCTCAGACAC ATACCAGGCC CAGGCCATGG TGGATATTGT CCGAGCCTTC AAGTGGAACT ATGTGTCCAC ACTGGCCTCA GAGGGCAGCT ACGGTGAGAG TGGTGTGGĀĞ 480 GCCTTTATCC AGAAGTCCCG AGAGAACGGA GGTGTGTGCA TTGCCCAGTC GGTGAAGATT CCACGGGAAC CCAAGACGGG GGAGTTCGAC AAGATCATCA AACGCCTACT GGAAACAŤCČ 600 AATGCCAGGG GTATCATCAT CTTTGCCAAC GAGGATGACA TCAGGAGGGT GTTGGAGĞČĀ GCTCGCAGGG CCAACCAGAC CGGCCACTTC TTTTGGATGG GTTCTGATAG CTGGGGCŤČČ AAGAGTGCCC CTGTGCTGCG CCTTGAGGAG GTGGCCGAGG GCGCAGTCAC CATTCTCCCC AAGAGGATGT CTGTTCGAGG GTTCGACCGA TACTTCTCCA GCCGCACGCT GGACAACAÃC AGGCGCAACA TCTGGTTTGC CGAGTTCTGG GAGGACAACT TCCATTGCAA GTTGAGCČGČ CACGCGCTCA AGAAGGGAAG CCACATCAAG AAGTGCACCA ACCGAGAGCG CATCGGGČĂĞ GACTCGGCCT ATGAGCAGGA GGGGAAGGTG CAGTTCGTGA TTGACGCTGT GTACGCCATG GGCCACGCGC TGCACGCCAT GCACCGTGAC CTGTGTCCCG GCCGCGTAGG ACTCTGCCCT 1080 CGCATGGACC CCGTGGATGG CACCCAGCTG CTTAAGTACA TCAGGAACGT CAACTTCTCA 1140 GGCATTGCGG GGAACCCTGT AACCTTCAAT GAGAACGGAG ACGCACCGGG GCGCTACGAC FIG. 9A.

31/32 ATCTACCAGT ACCAACTGCG CAATGGCTCG GCCGAGTACA AGGTCATCGG CTCGTGGACA

1260 GACCACCTGC ACCTCAGAAT AGAGCGGATG CAGTGGCCAG GGAGTGGCCA GCAGCTGCCG

1320 CGCTCCATCT GCAGTCTGCC CTGCCAGCCC GGGGAGCGAA AGAAGACTGT GAAGGGCĂTĞ

1380 GCTTGCTGCT GGCACTGCGA GCCCTGCACC GGGTACCAGT ACCAAGTGGA CCGCTACACC

TGTAAGACCT GCCCCTACGA CATGCGGCCC ACAGAGAACC GCACGAGCTG CCAGCCCATC

CCCATCGTCA AGTTGGAGTG GGACTCGCCG TGGGCCGTGC TGCCCCTCTT CCTGGCCGTG

GTGGGCATCG CCGCCACGCT GTTCGTGGTG GTCACGTTTG TGCGCTACAA CGATACCCCC

1620 ATCGTCAAGG CCTCGGGCCG GGAGCTGAGC TACGTGCTGC TGGCGGGCAT CTTTCTGTGC

1200

TACGCCACTA CCTTCCTCAT GATCGCAGAG CCGGACCTGG GGACCTGTTC GCTCCGCCGC

ATCTTCCTAG GGCTCGGCAT GAGCATCAGC TACGCGGCCC TGCTGACCAA GACCAACCGC

ATTTACCGCA TCTTTGAGCA GGGCAAACGG TCGGTCAGTG CCCCGCGTTT CATCAGCCCG

GCCTCGCAGC TGGCCATCAC CTTCATCCTC ATCTCCCTGC AGCTGCTCGG CATCTGCGTG

TGGTTCGTGG TGGACCCCTC CCACTCGGTG GTGGACTTCC AGGACCAACG GACACTTGAC

CCCCGCTTTG CCAGGGGCGT GCTCAAGTGC GACATCTCGG ACCTGTCCCT CATCTGCCTG

CTGGGCTACA GCATGCTGCT GATGGTCACG TGTACTGTGT ACGCCATCAA GACCCGAGGC

GTGCCGAGA CCTTCAACGA GGCCAAGCCC ATCGGCTTCA CCATGTACAC CACCTGCATT

2160

GTCTGGCTGG CCTTCATCCC CATCTTTTT GGCACCTCAC AGTCAGCCGA CAAGCTGTĂC

2220

ATCCAGACAA CCACACTGAC GGTCTCCGTG AGTCTGAGCG CTTCAGTGTC CCTGGGGATG

2280

CTCTACATGC CCAAAGTCTA CATCATCCTC TTCCATATTT TTCCATTCTG CTCCTGGCCT

FIG. 9B.

32/32

TCCCCTGCCA TCTGCCCTGC CCCCTGCCCC TCCTCCCTGA GCTGCCCCAT CCCCGCCATC

ATTITCTCTT CTGTTCCCCC TCGATCTCAT TTCCTACCAG CCTTCCCCCT ACTTGGCTTC

CTCCACCAAC TCTTTCACCA CGTTGC

FIG 9C

## INTERNATIONAL SEARCH REPORT

International Application No. PCT/US91/09422

		ON OF SUBJECT MATTER (if several		icate all) ³							
	_	ational Patent Classification (IPC) or to bo se See Attached Sheet.	th National Classification and IPC								
us cr	<b>.</b>	435/69.1, 240.2, 320.1; 53	0/350, 351, 387; 536/27.	B							
II. FIELI	DS SEAR		nentation Searched 4								
Classificati	ion System		Classification Symbols								
,	•.	US CL : 435/69.1, 240.2	1	197 - 536/27							
U.S.		03 (11 : 133/07:1, 210:2	, 320.1, 330/330, 331, 3	,0., 550,2,.							
	-0	Documentation Searched	other than Minimum Documentation	on.							
		•	ments are included in the Fields Se								
cas, c	online,	aps		*							
	•		*								
III. DOC	UMENTS	CONSIDERED TO BE RELEVANT 14									
Category*	<del>,</del>	n of Document, 16 with indication, where app	propriete, of the relevant passages 17	Relevant to Claim No. 18							
х/у	Nature et al inosit	e, Volume 325, issued 05 F "A new type of glutamatol phospholipid metabolismetire document.	ebruary 1987, Sugiyama	1-3, 6-8/9-30							
x/y	"Gluta major	n, Volume 3, issued July is mate receptor subtypes may catories: a study on Xer rat brain mRNA" pages 129 ent.	be classified into two popus occytes injected	1-3, 6-8/9-30							
<b>y</b> .,	et al.	e, Volume 342, issued 07 D , "Cloning by functional e glutamate receptor family atire document.	expression of a member	1-3 and 6-30							
x,p	al.,	e, Volume 349, issued 28 i "sequence and expression mate receptor", pages 760-7	n of a metabotropic	1-3, 6-30							
	<i>i</i> *										
"A" doc not "E" earli inte "L" doc or o doc or o doc but  IV. CER	not considered to be of particular relevance  E" sariier document but published on or after the international filing date  L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)  O" document referring to an oral disclosure, use, exhibition or other means  P" document referring to an oral disclosure, use, exhibition or other means  P" document published prior to the international filing date but later than the priority date claimed  V. CERTIFICATION  Date of the Actual Completion of the International Search ² Date of Mailing of this International Search Report ²										
Internatio	nal Search	ing Authority ¹	Signature of Authorized Officer	Manie !							
Te	A /TTS	•	Gian Wang Bh D 411	1							

#### FURTHER INFORMATION CONTINUED FROM PREVIOUS SHEETS (Not for publication)

I. CLASSIFICATION OF SUBJECT MATTER: IPC (5):

C12P 21/06; C12N 5/00, 15/00; C07H 15/12, 17/00; C07K 3/00, 13/00, 15/00, 17/00; A61K 35/14.

#### VI. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING This ISA found multiple inventions as follows:

#### Detailed reasons for holding lack of invention

The claims of the three groups have the characteristics of three distinct inventive concepts. Groups I-III are separate and distinct inventions, and require materially different considerations and searches.

#### Itemized summary of claims groupings

- I. Claims 1-3 and 6-30 are drawn to a method for producing a mammalian G protein by using its encoding sequence, classified in Class 435, subclass 69.1, 240.2; Class 530, subclass 387; Class 536, Subclass 27.

  II. Claims 4-5 and 31-33 are drawn to a method for determining the presence of a
- mammalian G protein by using monoclonal antibody, classified in Class 435, subclass 7,21; Class 424, subclass 85.8.
  III. Claims 34-38 are drawn to a method for identifying a compound, classified in
- Class 435, subclass 4.

• •		80	
-(vii)	IMMEDIATE SOURCE:		•
( , ,	(B) CLONE: ZC1752		
	PROBLEM SEC	TD NO:8:	
(xi)	SEQUENCE DESCRIPTION: SEC		
GATCCTTG1	AC AGAGCACAG		
(2) INFO	RMATION FOR SEQ ID NO:9:		
·	SEQUENCE CHARACTERISTICS		
(1)	(A) LENGTH: 22 base pair	rs	
	/pi mypr. nucleic acid		
	(C) STRANDEDNESS: SINGL		
	(D) TOPOLOGY: linear		
			,
(ii)	MOLECULE TYPE: CDNA		•
3			
	IMMEDIATE SOURCE:		*
(A11)	(B) CLONE: ZC2063		
,,			
		o TD NO.0.	
	SEQUENCE DESCRIPTION: SE	Q ID NO. 9	y .
	CT AGTAAAAGAG CT		6
	RMATION FOR SEQ ID NO:10:		•
(1)	SEQUENCE CHARACTERISTICS		
(-,	/X \ TENCTH: 14 Dase Das	rs	
-	(B) TVDE: NUCLEIC ECIC		
	(C) STRANDEDNESS: singl		.' '
_	(D) TOPOLOGY: linear		
	MOLECULE TYPE: CDNA		
(11)	, MOINCORD 2-1-		
. 1			
(vii)	IMMEDIATE SOURCE:		,
	(B) CLONE: ZC2064		
	<b>2</b> *		Ĭ
	) SEQUENCE DESCRIPTION: SI	EQ ID NO:10:	
(XI)	) SEQUENCE DESCRIPTION		
•	TAG TTTG		,
(2) THE	ORMATION FOR SEQ ID NO:11		
		·	
(i	) SEQUENCE CHARACTERISTIC	S:	_
	/XI TENGTH: 43 DASE PC	4	_
•	(B) TYPE: nucleic acid	le	
***	(C) STRANDEDNESS: sing (D) TOPOLOGY: linear		*
	(n) Totomer, Trues		

(ii) MOLECULE TYPE: CDNA

(vii)	IMME	STAIC	SC	OURCE:	
•	(B)	CLONE	: :	ZC2938	ì

(XI) SEQUENCE DESCRIPTION: SEQ ID NO:11:
GACAGAGCAC AGATTCACTA GTGAGCTCTT TTTTTTTTT TTT

43

- (2) INFORMATION FOR SEQ ID NO:12:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 21 base pairs
    - (B) TYPE: nucleic acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: CDNA
  - (Vii) IMMEDIATE SOURCE: (B) CLONE: ZC3015
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TTCCATGGCA CCGTCAAGGC T

21

- (2) INFORMATION FOR SEQ ID NO:13:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 21 base pairs
    - (B) TYPE: nucleic acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
    - (ii) MOLECULE TYPE: CDNA
    - (Vii) IMMEDIATE SOURCE: (B) CLONE: ZC3016
    - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AGTGATGGCA TGGACTGTGG T

21

- (2) INFORMATION FOR SEQ ID NO:14:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 20-base pairs
    - (B) TYPE: nucl ic acid
    - (C) STRANDEDNESS: singl
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: CDNA

•	,	* 5	•	· ·		, and the second se	•
·(vii)	TMMED	IATE SOUR	CE:	•			
(411)	/B):	CLONE: ZC	3652				
•		1			•		
,	à		, .				•
	CECTE	NOT DESCR	IPTION: SEC	ID NO:14:	· · · · · · · · · · · · · · · · · · ·		
(X1)	SEQUE	WCE DEDCA					20
	. mcc	meneral T	*				20
ACATGCACO	A TGC	TCTGIGI					* =
		W POD SEC	ID NO:15:		*		•
(2) INFOE	KMATIC	M FOR BEC			•		10
		nees attable	COPPLISTICS			•	
(1)	SEQUE	NCE CHAR	CTERISTICS:	rs	•	- x	
	(A)	LENGTH: 4	l base pail			•	•
•	(B)	TYPE: nuc	cleic acid	<b>.</b>	• • •		
· ·	(C)	STRANDED	TESS: single		·		•
	(D)	TOPOLOGY	TIMEAL	<b>∼</b> :.	(i) v		- 30
			mark	9			
(ii)	MOLEC	TULE TYPE	CUIA		•		
*			* · · · · · · · · · · · · · · · · · · ·		- a		* **
	•			• •	•		
(vii)	IMME	DIATE SOU	KCE:		1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	8.0	20
	(B)	CLONE: Z	C3654		•		
		• ,	ė.			•	• .
				O TO MO-15:			
(xi)	SEQUI	ence desci	RIPTION: SE	n in morta.			
•		,		110		2 *	21
AGTGATGG	CA TG	GACTGTGG '	r ·	191	· .		
				-0.0	Α		
(2) INFO	RMATI	on for se	Q ID NO:16:				
-		: .	0			V 10 m	
(i)	SEQU	ence char	ACTERISTICS				
	/ A Y	I.ENGTH:	5236 Dase P	airs	•	. 0	
,	(R)	TVPE: nu	cleic acid				
•	(C)	STRANDED	NESS: BINGT	e	**		*
•	(D)	TOPOLOGY	: linear				
	•						÷ '
(11)	MOLE	CULE TYPE	: CDNA			a	
(,,	•===						
• • •	•	i 1 2		•	- T		
(9711)	TMME	DIATE SOU	RCE:	*			•
( * * * * )	(B)	CLONE: S	N23				
	_,	ė.			· (0	•	•
/iv\	FEAT	URE:			*		
(14)	/A1	NAME/KEY	: CDS			1444	•
•	: (B)	LOCATION	1: 6273344	•			* * * * * * * * * * * * * * * * * * * *
· · · · · ,	(-/				*		•
			·			· ·	•
(41)	SEOU	ENCE DESC	RIPTION: SE	Q ID NO:16:			
(44)	2260					TACAGGGCAC	60
mx x < x x m	LV - detel	PAATACTC	TGGGAATTTT	ATTGGTGATG	CCTTTGTGTC	INCHOOCH	
TAAGAAT			•			Y COMMANDO A	120
1 common1		COTOTO	GTGAAGTGAT	GGGGGACTTG	TGGCTAGAGA	VACTITION	
ACGITCC	nun un			*		CAGAGCAGAG	180
		POTE GET CE	TGCTTGAGAG	AGGTCTGAGG	TTCTCAACAT	CHONDCHONG	, 200
TGGCCTI	MAN CI		- <del>-</del> ·			OTTOTATOGE	240
	733 <i>71</i>	- TOTAL CARA	TGCTAAGCCC	CCACTTCTCA	ACACTTAGTG	CICIGNICA	
CITICCAC	CAR G	**TT CUCUM				emements cec	300
		CACAACCG	CTGCAGTCCT	CTGACCTGAG	ACCAATAGCT	GIGICINCEC	
TGCCTGC	GAA C	-avanv-aa	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	· ·	•		

CURCULTITE SHEET

٠,			•						,	. 10	• "						**
GGÂ	CTCAG	CG T	CCAG	CTCA	c cgc	CAC	TAAC	GCG	CCCC	CA S	TTGG/	ACACO	CT G	ATCC!	CACA	,	360
CCT	TCGGG	CA C	CAGT	GAAA	A ACC	CGCGZ	ACTT	GAT	rttc:	rgg i	AAGA	ACGC	ec co	CAGGO	TGTG	٠.	420
GGA	GCGGT	CG T	GGAG(	GACC	A GC	AGGA	GGAA	GCG	GAGG	GGA (	GAGG	GGCA	GT AC	stgg?	AGGCA	· · .	480
•	AAAGC		_	•	•												540
	TCTGT																600
	CGGTG						እጥር	GTC	CGG	CTC	CTC	TTG	ATT	TTC	TTC	•	653
CCG	CGGTG	GA C	الالالالا	TCTT		Lnun	Met	Val	Arg	Leu	Leu	Leu	Ile	Phe	Phe		
						•	Ţ				•					·	
	ATG	300	mmm.	unic	GAG	ΔTG	тсс	ATT	TTG	CCC	AGG	ATG	CCT	GAC .	AGA		701
CCA	ATG Met	Tle	Dho	TAU	Glu :	Met	Ser:	Ile	Leu	Pro	Arg	Met	Pro	Asp .	Arg.		
10		116	F 1.0	:	15		· ·	•	,	20					25		•
_							· 		-1-6		maa	CTIC	aca	ACA	ATG	· · ·	749
λÄ	GTA	TTG	CTG	GCA	GGT	GCC	TCG	TCC	CAG	Z-GC	Cor	Val.	Ala	Ara	Met		
Lys	Val	Leu	Leu		GIA.	ATA.	Ser	Ser	38	ALG	261	142		40			
٠,		٠.		. 30	•	. •			٠						-		·
C3.6	GGA	CATT	CTC	ATC	ATC	GGA	GCC	CTC	TTC	TCA	GTC	CAT	CAC	CAG	CCT	* .	797
. GAC	GGA Gly	ASD	Val	Ile	Ile	Gly	Ala	Leu	Phe	Ser	Val	His	His	Gln	Pro		
			45					⊸ 50	<del>-</del> .	A	· ·	*	55	···			
			:							CCC	GAG	አጥሮ	AGG	GAA	CAG		845
CCI	A GCC	GAG	AAG	GTA	CCC	GAA	AGG	AAG	TGT	Glv	Glu	Ile	Ara	Glu	Gln		
Pro	A GCC		Lys	VAI	PTO	GIU	45 65	пåэ	Cys	0-1		70					
		60	:	•						٠	. •						
- ΤΔ'	r GGT	ATC	CAG	AGG	GTG	GAG	GCC	ATG	TTC	CAC	ACG	TTG	GAT	AAG	ATT		893
TV	r GGT r Gly	Ile	Gln	Arg	Val	Glu	Ala.	Met	Phe	His	Thr	Leu	Asp	гåа	TTG	:	*
	75		٠.	•		80		•		ē.	85	•				*.	•
	c GCG	· •				~	000	330	åπc	ACT	CTG	GGC	AGT	GAG	ATC	. *	941
AA	c GCG n Ala	GAC	CCG	GIG	TAU	Lau	Pro	Asn	Ile	Thr	Leu	Gly	Ser	Glu	Ile		
AS		Asp	PLO		95		•••	,	7	100	-	: _	1		105	Θ.	
	ī			•	_			•	•						CAA		989
CG	G GAC	TCC	TGC	TGG	CAC	TCT	TCA	GTG	GCT	CTC	GAA	CAG	AGC	TIA	Glu		JUJ
Ar	g gac g asp	Ser	Cys	Trp	His	Ser	Ser	Val	A18	Leu	GIU	GIU	Ser	120		• :	
	٠.		مد	. 110	•				. 113				•	· ·	•		
	C ATC	3.03	GAC	TCC	CTG	ATT	TCC	ATC	CGA	GAT	GAG	AAG	GAT	GGG	CTG		1037
. L.L	C ATC e Ile	AGA	Asp	Ser	Leu	Ile	Ser	Ile	Arg	Asp	Glu	Lys		Gly	Leu		
PII	e TTE	J	125		·,	4.		130				•	135				
	•					1	<b></b>					CCC	AGG	A CT	AAG	٠	1085
AA	C CGA	TGC	CTG	CCI	GAT	GGC	CAG	ACC	CIG	222	- Dev	GGC	Ara	Thr	Lys		
As	c cga n Arg	Cys	Leu	Pro	Asp	Gly	GTU	TIT	ren	FIU	FIO	150	J 3		1		
		140					145		. *				•	,	÷ ,		- سامريم د
	G CCI	) A MATE	الم	دعی	GTG	ATC	GGC	CCT	GGC	TCC	AGC	TCT	GTG	GCC	ATT	A	1133
A.A	G CCI	ATT	Ala	Glv	Val	Ile	Gly	Pro	Gly	Ser			Val	Ala	Ile		
тă	155	, 110		3		160	)		• =0		165	· .	•	• .	., ,		(2)
					· ·	. :== -				አመጣ		CAG	ATC	GCC	TAT		1181
C	A GT	CAG	AAT	CII	CTC	CAG	CIG	TTC	, GAC	Tla	Pr	Gln	Il	Ala	Tyr	•	
Gl	n Val	. Gln	Asn	Let	1 Leu 175	GIN	Leu	Live	, wah	180					185		
. 17	70	•			1/3				<del></del> -						•		

									84	, I				•			Ü.	
		•				0		-3-		3.⊖⊞	The state of	TAC	AAA	TAC.	TTC		1229	
TCI	: GCC	ACA	AGC	ATA	GAC	CTG I	AGT (	GAC	AAA	Thr	Leu	Tyr	LVS	Tyr	Phe	,	. * .	•
Ser	Ala	Thr	ser	IIe	Asp	Leu :	Ser .	Asp	195	,		-1-	-	200		•	•	•
,		• • •		190	e	•						•		;	•			
					m⊖ <del>in</del>	GAC	ACT	TTG	CAG	GCA	AGG	GCG	ATG	CTC	GAC		1,277	
CIG	AGG	GTG	GTC	Bro.	Sor	ASP	Thr	Leu	Gln	Ala	Arg	Ala	Met	Leu .	Asp			
Lev	Arg	Val	205	PLO.	J-C-2	•••		210	• • • •			,	215	· ·		· · · ·	* ***	
	·									:	003	CEC	CAC	ACA	GAA .	•	1325	
<b>Δ</b> Ψ2	GTC	AAG	CGT	TAC	AAC	TGG	ACC -	TAT	GTC	TCA	GCA	Up 1	Hic	Thr	Glu	*		
Ile	Val	Lys	Arg	Tyr	Asn	Ττħ	T17T	-1-	Val	Ser	WTG	230	41 T G					· :
	• -0-	.220			. 8	•					8.			, .		•	•	•
					3.4M	GGA	አጥር	CAT	GCT	TTC	AAA	GAA	CTG	GCT	GCC		1373	
GG	ea e	TAC	GGC	GAG	AGT	GGA Gly	Met	ASD	Ala	Phe	Lys	Glu	Leu	Ala	Ala.	•		
Gly	y Asr	Tyr	GIA	GIU	Ser	240	••••				245			•		•		
	235	).	•	•					•			m> 4	3.00	 3.3.TD	COT	.:	1421	
CÀ	c car	GGC	CTC	TGC	ATC	GCA	CAC	TCG	GAC	AAA	ATC	TAC	AGC	Vel .	Ala	:		
Gli	n Clu	Gly	Leu	Cys	Ile	GCA Ala	His	Ser	ASP	Lys	TIE	TYL	Jer		265	,		
25	0			r	255								•					
	• :				43.0	CGG		CTG	CGT	AAA	CTC	CGG	GAG	CGG	CTT		1469	
GG	C GAG	AAG	AGC	LIL	GAC	Arg	Teu	Leu	Arq	Lys	Leu	Arg	Glu	Arg	Leu		1	
GI	y Glu	TAS	Ser	270	Yob	<b>11.</b> 3			275	<u>.</u>	•			280	,			,
				_		. 0	··			٠	· ·			-cmc	cċc		1517	
-	CAA	: GCC	AGG	GTT	GTG	GTC	TGC	TTC	TGC	GAG	GGC	ATG	ACA The	(Ia)	Arg	• .		. 2
Pr	o LV	s Ala	Àrġ	Val	Val	Val	Cys	F 110		Glu	GIY	met	295	Val	444.3			
			285					450	. 6	. 0						` -	,	
		•	•			CGC	CCC	CTC.	GGC	GTC	GTG	GGC	GAG	TTC	TCA		1565	
GG	CTT	A CTO	AGT	GCC	ATG	CGC	Ara	Leu	Gly	Val	Val	Gly	Glu	Phe	Ser		,	
Gl	y Le	ı Let	1 Ser	, VTG	Mec	wra	305				,	310			•			٠
		300		•				8		·			100	CNA	GGC		1613	
CTI	יינג יי	r GG)	AGT	GAT	GGA	TGG	GCA	GAC	AGA	GAT	GAA	GTC	TIA	Glu	Glv	-		
I.e	u Il	e Gly	Ser	Asp	Gly	TGG	Ala	Asp	Arg	Asp	.22£	AGT	116		<u></u>			
	31	5		• •	•	320					•		7			3		
		٠,				GGA	000	አጥር	ACA	ATA	ÄAG	CTT	CAG	TCT	CCA		1661	
TA	T GA	g GI(	GAA	CGCC	AAC	GGA	Glv	Ile	Thr	Ile	Lys	Leu	Gln	Ser	Pro	-	•	
		u Va.	r Gra	i VII	335	GIY		,		340	),				345			
33	0			•				•						- C1C	»CC		1709	• 1
G B	.a. at	c AG	G TC	· TTI	GAT	GAC	TAC	TTC	CTG	AAG	CIG	AGG	CIG	ARD	Thr	•		
G]	u Va	1 Ar	g Sei	Phe	AST	Asp	Tyr	Phe			Leu	MLG	, 555	360	;		*	
			•	356	,													
	•	,				TTC	COT	GAG	TTC	TGG	CAA	CAT	CGC	TTC	CAG	•	1757	
A	AC AC	A AG	G AAT		TGG	Phe	Pro	Glu	Phe	TI	Gln	His	Arg	Phe	Gln			
As	sn Th	r Ar	g ASI 36!	i Pro	) TTP	Pile		370	)	_	•		375		,	•		
	· ·									-					GTG.		1805	
me	:m : 00	C CT	A CC	r GG	A CAC	CTC	TTG	GA	AA(	c ccc	AAC	TT	L' AAG	T.VO	Val			
	sa De	a Le	u Pro	G1	y His	c Crc			ı Ası	n Pro	ASI	1 LUG	U ≥ nas					<u>:</u> .
		38	0				300											•
		•	•	<u> </u>		C TTG	۳.۶ س	ca:	<b>.</b>	C TAT	r GTC	CAC	G GAC	AGC	: AAA		1853	
T	GC AC	A GG	A AA	r gai	A AG	r Lev	GAA	. GA	u Asi	n Ty	r Val	L Gli	n Ası	Ser	Lys			•
C	ys Tì	r Gl	y As	u GT	u 26;	400		- <del>-</del>		• "	405	5						
	. 39	5		•						_		1			i Cac		1901	
` **	TC C	्र । ।	ጥ Gጥ	C AT	C AA	r GCC	: ATC	TA'	T GC	C AT	G GCI	A CA	ı GG(	e CIU		' س	,	
A.		-	. <del></del>			-30								QI	IRST	TÚTI	e she	ET
	•					•	.*							J.	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,			

ANC ATG CAC CAT GCT CTG TGT CCC GGC CAT GTG GGC CTG TGT GAT GCT  ANS MET HIS HIS Ala Leu Cys Pro Gly His Val Gly Leu Cys Asp Ala  ATG AAA CCC ATT GAT GGC AGG AGG AGG CTC CTG GAT TTC CTC ATC AAA TCC  MET Lys Pro 11e Asp Gly Arg Lys Leu Leu Asp Phe Leu Ile Lys Ser  445  TCT TTT GTC GGA GTG TCT GGA GAG GAG GTG TGT TTC GAT GAA AAA TCC  Ser Phe Val Gly Val Ser Gly Glu Glu Val TTP Phe Asp Glu Lys Gly  465  GAT GCT CCC GGA AGG TAT GAC ATT ATG AAT CTG CAG TAC ACA GAA GCT  ASP Ala Pro Gly Arg Tyr Asp Ile Met Asn Leu Gln Tyr Thr Glu Ala  A75  AAT CGC TAT GAC TAT GTC CAC GTG GGG ACC TGG CAT GAA GAG GTG CTG  AAN ATT GAT GAC TAT GTC CAC GTG GGG ACC TGG CAT GAA GAG GTG CTG  AAN ATT GAT GAT TAC AAA ATC CAG ATG AAA AAA ACC GGA ATC GTA CGA  ASP Alle Asp Asp Tyr Lys Ile Gln Met Asn Lys Ser Gly Met Val Arg  510  TCT GTG TGC AGT GAC CTT TGC TTA AAG GAT CAG ATT ATG GAT CTA CGA  ASP Ile Asp Asp Tyr Lys Ile Gln Met Asn Lys Ser Gly Met Val Arg  510  TCT GTG TGC AGT GAC TGC TGC TTA AAG GAT CAG ATT ATG CGA ATT ATG GAT CTA CGA CTA CGA ATT ATG GTG TGC TAT CAG CYS Leu Lys Gly Gln Ile Lys Val Ile Arg  520  AAA GGA GAA GTG AGC TGC TGC TTA AAG GTC CAG CTC TGC ATA CGG  AAA GGA GAA GTG AGC TGC TGC TTA AAG CGC TGC AGA GCC TGC GAC TTG CAC TGC TGC TGC TGC TGC TGC TGC TGC TGC TG			Gly	Phe	Val	Ile	Asn 415	Ala	Ile	Tyr	Ala	Met	Ala	His	Gly	Leu	Gln 425		ž.
ASN MET HIS HIS Als Leu Cys Pro Gly HIS 435  ATO 440  ATO 440  ATO 4415  ATO			; 				•	mem.		CCC	ርኔጥ '	:	GGC	CTG	TGT	GAT	GCT	· 1	.949
Met Lys Pro 11e Asp Gly Arg Lys Leu Asp File 2445         455           TCT TTT GTC GGA GTG TCT GGA GAG GAG GTG TGG TTC GAT GAG AAG GGG AGG GTG TGT TTT GTC GGA AGG GTG TAT GAT ATT ATT ATT ATT TTT PHE ASP GLU Lys GLY 465         2045           GAT GCT CCC GGA AGG TAT GAC ATT ATG AAT CTG CAG TAC ACA GAA GCT ASP ALA Pro Gly Arg Tyr Asp Ile Met Asn Leu Gln Tyr Thr Glu Ala ASP ALA Pro Gly Arg Tyr Asp Ile Met Asn Leu Gln Tyr Thr Glu Ala ASP ASP Tyr Lys Lie Glo Trop His Glu Gly Val Leu Gly Thr Trp His Glu Gly Val Leu Gly Asn Ile Asp Asp Tyr Lys Ile Gln Met Asn Lys Ser Gly Met Val Arg Sol TCT GTG TGC AGT GAG CCT TGC TTA AAG GGT CAG ATT AAG GTC ATA CGG CTG TGT GTG TGC AGT GAG CCT TGC TTA AAG GGT CAG ATT AAG GTC ATA CGG CTG TGC TGC TGC TGC TGC AGA GCC TGC AAA GAC AAT Sol Tle Arg Sol Tyr Glu Val Ser Cys Cys Trp Ile Cys Thr Ala Cys Lys Glu Asn Sol Tyr Glu Val Ser Cys Cys Trp Ile Cys Thr Ala Cys Lys Glu Asn Sol Tyr Glu Phe Val Gln Asp Glu Phe Thr Cys Arg Ala Cys Asp Leu Gly Trp Sol Trp Pro Asn Ala-Glu Leu Thr Gly Cys Glu Pro Ile Pro Val Arg Tyr Tyr For Asn Ala-Glu Leu Thr Gly Cys Glu Pro Ile Pro Val Arg Tyr Tyr For Asn Ala-Glu Leu Thr Gly Cys Glu Pro Ile Pro Val Arg Tyr Tyr Asp Ile Glu Ser Ile Ile Ala Ile Ala Phe Ser Cys Goo CTG GGC ATC CTC GTG ACA CTC ACC CTC ATC TTC CTT CTC TGC CTT TCT TGC GGC ACC CTC GGC ACC CTC GTG ACC CTC TTT CTT TGC GTG GGC ACC ACC GCC ATC CTC GTG ACC CTC ACA CTC ACC CTC ATC TTC CTT CTC TAC CTC GGC GAC ACC CTC GTG ACC CTC ATC TTC T		AAC Asn	ATG Met	His	CAT His	Ala	Leu	Cys	Pro	Gly	HIS	Val	Gly	Leu	Cys	vah	Ala	•	
Ser Phe Val Gly Val Ser Gly Glu Glu Val Trp He Asp CT Sys CT 460  GAT GCT CCC GGA AGG TAT GAC ATT ATG AAT CTG CAG TAC ACA GAA GCT ASP Ala Pro Gly Arg Tyr Asp Ile Met Asn Leu Gln Tyr Thr Glu Ala 485  AAT CGC TAT GAC TAT GTC CAC GTG GGG ACC TGG CAT GAA GGA GTG CTG 495  AAT ATT GAT GAT TAC AAA ATC CAG ATG AAC AAA AGC GGA ATG GTA CGA ASP Ala Pro Gro Cag Gu Pro Cys Leu Lys Gly Gln Ile Lys Val Ile Arg 510  TCT GTG TGC AGT GAG CCT TGC TTA AAG GGT CAG ATT AAG GTC ATA CGG 2237  AAA GGA GAA GTG AGC TGC TGC TGA ATC ACG GCC TGC AAA GAG AAT ASP 525  AAA GGA GAA GTG AGC TGC TGC TGA ATC TGC ACG GCC TGC AAA GAG AAT ASP 525  AAA GGA GAA GTG AGC TGC TGC TGG ATC TGC ACG GCC TGC AAA GAG AAT 550  GAG TTT GTG CAG GAC GAC GAC TTC ACC TGC AGA GCC TGT GAC CTG GGG TGG CCC TGA AAC GAC ACA GGC TGT GAC CTG GGC TGC AAC GCC TGT GAC CTG GGG TGG CCC AAC GCC GCC AAC GCC TGT GAC CTG GGC TGC ACC GCC TGC AAC GCC TGT GAC CTG GGG TGG CCC TGC AAC GCC TGT GAC CTG GGC TGC ACC GCC TGC AAC GCC TGT GAC CTG GGC TGC ACC GCC TGC AAC GCC TGT GAC CTG GGC TGC ACC GCC TGC AAC GCC TGT GAC CTG GGC TGC GCC TGC AAC GCC TGC AAC GCC TGT GAC CTG GGC TGC GCC TGC AAC GCC TGC AAC GCC TGT GAC CTC GCC TATC TCC ACC GCC TATC CTC TGC ACC GCC TGT GAC CTC GTC CGT TAT TCC GCC TGT GAC CTC TGT GAC CTC TGC ACC GCC TGT GAC CTC TGC TTC TGC TGC GCC TATC TCC GTG ACC CTC TGT ACC CTC TGT ACC CTC TGC ACC CTC TGT TGC ACC CTC TGT TGT TGT TGT ACC CTC TGT TGT TGT ACC CTC TGT TGT TGT TGT TGC CTC TGT TGT T		ATG Met	AAA Lys	CCC	Ile	GAT Asp	GGC Gly	AGG Arg	AAG Lys	Leu	CTG Leu	GAT Asp	TTC Phe	CTC Leu	776	AAA Lys	TCC Ser	1	997
AND ALB PRO GLY ARG TYP ASP ILE MET ASH LEW SITE 14 AND ARE		TCT Ser	TTT Phe	Val	GGA Gly	GTG Val	TCT Ser	GGA Gly	Glu	GAG Glu	GTG Val	TGG Trp	Lue	vah	GAG Glu	AAG Lys	GGG Gly	· · 2	2045
ASR Arg Tyr Asp Tyr Val His Val Gly Thr Tip his Glo Gly Tyr Sobs  AAT ATT GAT GAT TAC AAA ATC CAG ATG AAC AAA AGC GGA ATG GTA CGA SID Sobs Sobs Sobs Sobs Sobs Sobs Sobs Sobs		GAT Asp	Ala	Pro	GGA Gly	AGG Arg	TAT Tyr	Asp	ATT	ATG Met	AAT Asn	CTG Leu	GTII	TAC	ACA Thr	GAA Glu	GCT Ala	2	2093
ASR I le ASP ASP TYP Lys I le Gin Met Ash Lys Set Gif Met Ash Lys Gif Met Ash Lys Gif Met Ash Lys Gif Met Ash Ged Att Acc Gif Att Acc Cys Set Gif Met Ash Ged Att Lys Gif Met Ash Gif Acc Cys Lys Gif Met Ash Ged Att Lys Gif Gif Gif Acc Acc Cys Cys Trp I le Cys Thr Ala Cys Lys Gif Ash Cys Lys Gif Ash Cys Asp Leu Gif Trp Set Acc Cys Acc Gif Ged Cys Acc Cys Asp Leu Gif Trp Cys Arg Ala Cys Asp Leu Gif Trp Cys Ash Ala Gif Acc Cys Gif Cys Gif Pro I le Pro Val Arg Tyr Set Gif Gif Gif Cys Gif Gif Gif Cys Gif Gif Gif Cys Cys Cys Gif Cys Cys Cys Gif Cys		Asn	CGC	TAT	GAC Asp	TAT Tyr	Val	CAC His	GTG Val	GGG Gly	ACC	Trp	CAT	GAA Glu	GGA Gly	GTG Val	100		2141
Ser Val Cys Ser Glu Pro Cys Leu Lys Gly Gln lie Lys 535  AAA GGA GAA GTG AGC TGC TGC TGG ATC TGC ACG GCC TGC AAA GAG AAT 2285 Lys Gly Glu Val Ser Cys Cys Trp lie Cys Thr Ala Cys Lys Glu Asn 550  GAG TTT GTG CAG GAC GAG TTC ACC TGC AGA GCC TGT GAC CTG GGG TGG 2333 Glu Phe Val Gln Asp Glu Phe Thr Cys Arg Ala Cys Asp Leu Gly Trp 555  TGG CCC AAC GCA GAG CTC ACA GGC TGT GAG CCC ATT CCT GTC CGT TAT 2381 Trp Pro Asn Ala-Glu Leu Thr Gly Cys Glu Pro 11e Pro Val Arg Tyr 570  CTT GAG TGG AGT GAC ATA GAA TCT ATC ATA GCC ATC GCC TTT TCT TGC 590  CTG GGC ATC CTC GTG ACG CTG TTT GTC ACC CTC ATC TCT GTC CTG TAC 12475 Leu Glu Trp Ser Asp 11e Glu Ser 11e 11e Ala 11e Ala Phe Ser Cys 600  CTG GGC ATC CTC GTG ACG CTG TTT GTC ACC CTC ATC TTC GTT CTG TAC 12475 Leu Gly Il Leu Val Thr Leu Phe Val Thr Leu 11e Phe Val Leu Tyr 605  CGG GAC ACA CCC GTG GTC AAA TCC TCC AGT AGG GAG CTC TGC TAT ATC 2521 Arg Asp Thr Pr Val Val Lys Ser Ser S r Arg Glu Leu Cys Tyr 11e 620  CTG GCC ACC CTC ATC CTC ACC CTC TTC ACC CTC TTC ACC CTC ATC TCC ACC CTC ATC TCC TGC TAT ATC 2521  CGG GAC ACA CCC GTG GTC AAA TCC TCC AGT AGG GAG CTC TGC TAT ATC 2521  Arg Asp Thr Pr Val Val Lys Ser Ser S r Arg Glu Leu Cys Tyr Ile 620		AAT Asn	ATT Ile	GAT Asp	GAT Asp	Tyr	Ļys	ATC Ile	CAG Gln	ATG Met	ASI	AAA Lys	AGC Ser	GGA Gly	ATG Met	V 44.4	CGA Arg	·	2189
Lys Gly Glu Val Ser Cys Cys TTP 11e Cys Int ACC 550  GAG TTT GTG CAG GAC GAG TTC ACC TGC AGA GCC TGT GAC CTG GGG TGG Glu Phe Val Gln Asp Glu Phe Thr Cys Arg Ala Cys Asp Leu Gly Trp 555  TGG CCC AAC GCA GAG CTC ACA GGC TGT GAG CCC ATT CCT GTC CGT TAT TTP Pro Asn Ala Glu Leu Thr Gly Cys Glu Pro 11e Pro Val Arg Tyr 570  CTT GAG TGG AGT GAC ATA GAA TCT ATC ATA GCC ATC GCC TTT TCT TGC Leu Glu Trp Ser Asp 11e Glu Ser 11e 11e Ala 11e Ala Phe Ser Cys 600  CTG GGC ATC CTC GTG ACG CTG TTT GTC ACC CTC ATC TTC GTT CTG TAC Leu Gly Il Leu Val Thr Leu Phe Val Thr Leu Ile Phe Val Leu Tyr 605  CGG GAC ACA CCC GTG GTC AAA TCC TCC AGT AGG GAG CTC TGC TAT ATC ATG ASP Thr Pr Val Val Lys Ser Ser S r Arg Glu Leu Cys Tyr Ile 620  CTG GGC ATC CTC ATC CTC ATC TTC ACC CTC ATC TTC III CTC ATG CGC TAT CTG TGC CCT TTC ACC CTC ATC CGG GAC ACA CCC GTG GTC AAA TCC TCC AGT AGG GAG CTC TGC TAT ATC ATG ASP Thr Pr Val Val Lys Ser Ser S r Arg Glu Leu Cys Tyr Ile 620  CTG GGC ATC CTC ATC CT		TCT Ser	GTG Val	TGC Cys	Ser	Glu	CCT	TGC	TTA	Lys	GTÅ	CAG Gln	ATT	AAG Lys	AGT	ATA Ile	CGG		2237
Glu Phe Val Gln Asp Glu Phe Thr Cys Arg Ala Cys Asp Data Cys Sep 555  TGG CCC AAC GCA GAG CTC ACA GGC TGT GAG CCC ATT CCT GTC CGT TAT 2381  Trp Pro Asn Ala Glu Leu Thr Gly Cys Glu Pro Ile Pro Val Arg Tyr 580  CTT GAG TGG AGT GAC ATA GAA TCT ATC ATA GCC ATC GCC TTT TCT TGC 2429  Leu Glu Trp Ser Asp Ile Glu Ser Ile Ile Ala Ile Ala Phe Ser Cys 600  CTG GGC ATC CTC GTG ACG CTG TTT GTC ACC CTC ATC TTC GTT CTG TAC Leu Gly Il Leu Val Thr Leu Phe Val Thr Leu Ile Phe Val Leu Tyr 605  CGG GAC ACA CCC GTG GTC AAA TCC TCC AGT AGG GAG CTC TGC TAT ATC ATG ASP Thr Pr Val Val Lys Ser Ser S r Arg Glu Leu Cys Tyr Ile 620  CTG GGC ATC CTC GTG GTC CGC TAT GTG TGC CCT TTC ACC CTC ATC 257		AAA Lys	GGA Gly	Glu	Val	AGC Ser	TGC	TGC Cys	Trp	TTE	TGC Cys	ACG	GCC Ala	CJO	-3-	GAG Glu	AAT Asn	*	2285
TGG CCC AAC GCA GAG CTC ACA GGC TGT GAG CCC ATT CCT GTC CGT TAT Trp Pro Asn Ala Glu Leu Thr Gly Cys Glu Pro Ile Pro Val Arg Tyr 570  CTT GAG TGG AGT GAC ATA GAA TCT ATC ATA GCC ATC GCC TTT TCT TGC Leu Glu Trp Ser Asp Ile Glu Ser Ile Ile Ala Ile Ala Phe Ser Cys 590  CTG GGC ATC CTC GTG ACG CTG TTT GTC ACC CTC ATC TTC GTT CTG TAC Leu Gly Il Leu Val Thr Leu Phe Val Thr Leu Ile Phe Val Leu Tyr 605  CGG GAC ACA CCC GTG GTC AAA TCC TCC AGT AGG GAG CTC TGC TAT ATC Arg Asp Thr Pr Val Val Lys Ser Ser S r Arg Glu Leu Cys Tyr Ile 620  CGC GAC ACA CCC GTG GTC CTC GCC TAT GTG TGC CCT TTC ACC CTC ATC 625		GAG Glu	Phe	. Val	CAG Gln	GAC Asp	GAG Glu	Phe	Thr	TGC Cys	AGA Arg	GCC Ala	Cys	Voh	CTG Leu	GGG Gly	TGG	*	2333
CTT GAG TGG AGT GAC ATA GAA TCT ATC ATA GCC ATC GCC ATC GCC CYS  Leu Glu Trp Ser Asp Ile Glu Ser Ile Ile Ala Ile Ala Phe Ser Cys  590  CTG GGC ATC CTC GTG ACG CTG TTT GTC ACC CTC ATC TTC GTT CTG TAC  Leu Gly Il Leu Val Thr Leu Phe Val Thr Leu Ile Phe Val Leu Tyr  605  CGG GAC ACA CCC GTG GTC AAA TCC TCC AGT AGG GAG CTC TGC TAT ATC  Arg Asp Thr Pr Val Val Lys Ser Ser S r Arg Glu Leu Cys Tyr Ile  625  626  630  CTG GTC ATC CTC ATC CTC ACC CTC ATC A		Trp	CCC		GCA Ala	GAG - Glu	Let	1 Thr	GGC	TGI Cys	GAG Glu	PIO	77.0	CCI Pro	GTC Val	CGT	' ISI		2381
CGG GAC ACA CCC GTG GTC AAA TCC TCC AGT AGG GAG CTC TGC TAT ATC Arg Asp Thr Pr Val Val Lys Ser Ser S r Arg Glu Leu Cys Tyr Ile 620 625 626 627 627 628 627 628 628 629 627 628 628 629 620 625 627 627 628 627 628 627 628 627 628 628 628 628 628 628 628 628 628 628		CII	GAC	TGG Tr	AGT Ser	Asp	) Ile	A GAA e Glu	TCI Ser	T ATO	3 TTE	NTG	ATC	GCC Ala	TTT	501	, -,-		2429
CGG GAC ACA CCC GTG GTC AAA TCC TCC AGT AGG GAG CIC ISO INTO ILE Arg Asp Thr Pr Val Val Lys Ser Ser S r Arg Glu Leu Cys Tyr Ile 630 625 630	•	CTG	GG(	C ATO	Leu	l Val	ACC L Thi	CTG Leu	TTI Phe	3 AST	LTIN	CTC Leu	ATC Ile	TTC Phe	,		TAC		2477
257		CGG	GA(	p Thi	Pr	C GTC Val	GTO	C AAA l Lys	s Sei	r Se	C AGI	r Agg	GAG	The s	. Cyc	TAT	r ATC	*	2525
		ATT	CTC			r AT	r TT(	c cro	GGG GGI	C TA'	r GTO	G TGC	CCI Pro	r TT(	C ACC	CTC	C ATC		2573

	:					-40					645					10
	635			•		640				•						2621
GCC Ala	AAA Lys	CCT Pro	ACT Thr	ACC Thr	Thr	TCC Ser	TGC Cys	TAC Tyr	nea	CAG Gln 660	CGC	CTC	CTA Leu	GTT Val	GGC Gly 665	2621
650	÷				625	<b></b>	mom.	COTT	ביואיוי	GTG	ACC	AAA	ACC	AAT	CGT	2669
CTC	TCT Ser	TCT Ser	Ala	Met 670	Cys	Tyr	Ser	WTO	Leu 675	Val	Thr	Lys	Thr	Asn 680	Arg	-
ATT	GCA Ala	CGC Arg	ATC Ile 685	CTG Leu	GCT Ala	GGC Gly	AGC Ser	AAG Lys 690	AAG Lys	AAG Lys	ATC Ile	TGC Cys	ACC Thr 695	CGG Arg	AAG Lys	2717
CCC Pro	AGA Arg	TTC Phe 700	ATG Met	AGC Ser	GCT Ala	TGG Trp	GCC Ala 705	CAA Gln	GTG Val	ATC Ile	ATA Ile	GCC Ala 710	TCC	ATT Ile	CTG Leu	2765
ATT Ile	AGT Ser 715	GTA Val	CAG Gln	CTA	ACA Thr	CTA Leu 720	GTG Val	GTG Val	ACC Thr	TTG Leu	ATC Ile 725	ATC	ATG Met	GAG Glu	CCT Pro	2813
CCC Pro	ATG Met	CCC Pro	ATT Ile	TTG Leu	TCC Ser 735	TAC Tyr	CCG	AGT Ser	ATC Ile	AAG Lys 740	GAA Glu	GTC Val	TAC Tyr	CTT Leu	ATC Ile 745	2861
	AAT Asn	ACC	AGC Ser	AAC Asn 750	Leu	GGT Gly	GTA Val	AGT	GCC Ala 755	CCT Pro	GTG Val	GGT Gly	TAC Tyr	AAT Asn 760	GGA Gly	2909
CTC Leu	CTC Leu	ATC Ile	ATG Met 765	Ser	TGT Cys	ACC	TAC	TAT TYP 770	GCC Ala	TTC Phe	AAG Lys	ACC	CGC Arg 775	AAC Asn	GTG Val	2957
CCG	GCC Ala	AAC Asn 780	Phe	AAT Asn	GAG Glu	GCT	AAA Lys 785	TAT	ATC Ile	GCC Ala	TTC Phe	ACC Thr 790	ATG Met	TAC	ACT Thr	3005
ACC Thi	TGC Cys	Ile	ATC Ile	TGG Trp	CTG Leu	GCT Ala 800	LITE	GTT Val	CCC	ATT	TAC Tyr 805	TTT	GGG	AGC Ser	AAC Asn	3053
TAC Ty:	: Lys	ATC	ATC	ACT	ACC Thr 815	Cys	TTC	GCG Ala	GTG Val	AGC Ser 820		AGT	GTG Val	ACG	Val 825	3101
	•	GGG Gly	TGC	ATG Met	Pne	ACT Thr	CCG	AAG Lys	ATG Met 835	-1-	ATC	ATC	ATT	GCC Ala 840	AAA Lys	3149
CC! Pro	GAG Glu	AGG	AAC Asn 845	Agi	CGC Arg	AGT Ser	GCC	TTC Phe 850		ACC	TCT	GAT Asp	GTT Val 855	GTC Val	CGC	3197
AT(	CAC His	GTC Val	GGI Gly		GGC Gly	: AAA Lys	CTG Lev 865	FL	TGC Cys	CGC Arg	TCC Ser	AAC ABT	ACC Thr	TTC Phe	CTC Leu	3245

AAC ATT TTC CGG AGA AAG AAG CCC GGG GCA GGG AAT GCC AAG AAG AGG	3293
AAC ATT TTC CGG AGA AAG AAG CCC GGG GGA GAN ALA Lys Lys Arg Asn Ile Phe Arg Arg Lys Lys Pro Gly Ala Gly Asn Ala Lys Lys Arg 880 885	
CAG CCA GAA TTC TCG CCC AGC AGC CAG TGT CCG TCG GCA CAT GCG CAG	3341
cin pro Giu phe Ser Pro Ser Ser Gin Cys Pro Ser Ala	- 10
890 895	
CTT TGAAAACCCC CACACTGCAG TGAATGTTTC TAACGGCAAG TCTGTGTCAT	3394
Leu	
GGTCTGAACC AGGTGGAAGA CAGGCGCCCA AGGGACAGCA CGTGTGGCAG CGCCTCTCTG	3454
TGCACGTGAA GACCAACGAG ACGGCCTGTA ACCAAACAGC CGTAATCAAA CCCCTCACTA	3514
AAAGTTACCA AGGCTCTGGC AAGAGCCTGA CCTTTTCAGA TGCCAGCACC AAGACCCTTT	3574
ACANTGTGGA AGAAGAGGAC AATACCCCTT CTGCTCACTT CAGCCCTCCC AGCAGCCCTT	3634
CTATGGTGGT GCACCGACGC GGGCCACCGG TGGCCACCAC ACCACCTCTG CCACCCCATC	3694
TGACCGCAGA AGAGACCCCC CTGTTCCTGG CTGATTCCGT CATCCCCAAG GGCTTGCCTC	3754
CTCCTCTCCC GCAGCAGCAG CCACAGCAGC CGCCCCCTCA GCAGCCCCCG CAGCAGCCCA	3814
AGTCCCTGAT GGACCAGCTG CAAGGCGTAG TCACCAACTT CGGTTCGGGG ATTCCAGATT	3874
TCCATGCGGT GCTGGCAGGC CCGGGGACAC CAGGAAACAG CCTGCGCTCT CTGTACCCGC	3934
CCCCCCTCC GCCGCAACAC CTGCAGATGC TGCCCCTGCA CCTGAGCACC TTCCAGGAGG	3994
CCCCGCCTCC GCCGCAACAC CIGCAGAIGC IGCCCCTGC.	4054
AGTCCATCTC CCCTCCTGGG GAGGACATCG ATGATGACAG TGAGAGATTC AAGCTCCTGC	4114
AGGAGTTCGT GTACGAGCGC GAAGGGAACA CCGAAGAAGA TGAATTGGAA GAGGAGGAGG	
ACCTGCCCAC AGCCAGCAAG CTGACCCCTG AGGATTCTCC TGCCCTGACG CCTCCTTCTC	4174
CTTTCCGAGA TTCCGTGGCC TCTGGCAGCT CAGTGCCCAG TTCCCCCGTA TCTGAGTCGG	4234
TCCTCTGCAC CCCTCCAAAT GTAACCTACG CCTCTGTCAT TCTGAGGGAC TACAAGCAAA	4294
GCTCTTCCAC CCTGTAGTGT GTGTGTGTGT GTGGGGGGGGGG	4354
CCAGAGATGC CAAGGAGTGT CAACCCTTCC AGAAATGTGT AGAAAGCAGG GTGAGGGATG	4414
GGGATGGAGG ACCACGGTCT GCAGGGAAGA AAAAAAAAA TGCTGCGGCT GCCTTAAAGA	4474
AGGAGAGGA CGATGCCAAC TGAACAGTGG TCCTGGCCAG GATTGTGACT CTTGAATTAT	4534
TCAAAAACCT TCTCTAGAAA GAAAGGGAAT TATGACAAAG CACAATTCCA TATGGTATGT	4594
AACTTTTATC GAAAAAATA ATAAAACGTA AAAATAAAAT	4654
AACTITIATC GAAAAAAIA AIAISAGCII COCACTCCCG TGGTAAAACT AGAAGCGAAG	4714
TTTGCTCAAT CGTGCATACA TATAICIGCC CACATOTOO	477
CAGGCCCTGC GATGGTGCCA ACTGAATCCT AAGTTCATCT	r .

	•		-					8	8 /	100		t t					•
CNCGC	:CNG0	ià GC	CGAC	AGG	CAC	GAGG	CGG	GGGT	AGGI	TC G	GAC	ACAG	C TO	CCAT	CTCA	.4	834
CACCO	PTC & C	אוי פי	rgete	AGT	C TT	CAGAC	TCC	TGG	CTA	AGG A	AGAC	CCGG	G G	CTG	ACCTT	. 4	894
ATCA(	ccett Conc	re e	TTTC	ĈACTO	CT	GTGAT	CCA	TTG	CCAGO	CT C	TAG	CAC	c GC	GAT	AAAGG	. 4	954
CACAC	344TE	C T	TTTG	CATT	CTC	GTGAT	TCC	CTG	rg <b>t</b> ti	raa (	GAA	\AGG!	A ÀC	STAT	GAGCA	5	5014
AAGC	PATC	AC C	AAAA	AGAG	C GC	CATTA	AGAA	GTT	ACGG	GGG 1	AGAAI	AAAA	AG AG	GAAG	CAAGA		5074
тсат:	TAT	AA G	CACA	GGC	C TT	GAAC	AAGG	TGA	GCGT	GCT [	rcac:	AGAT.	rc co	GTAT	TAATG	,	5134
TACA	GATA(	CT T	TTGG	AGAG	G AG	aaag/	ATAA	CAA	GGAG'	rgt (	CAGG	CCGT	rt G	rgaa	CTCAC		5194
TTGC																	5236
(2)	NFO	RMAT	ION	FOR	SEQ	ID NO	0:17	<b>.</b> 7,		•	٠	•	8	*			
*			EQUE (A)	NCE LEN TYP	CHAR GTH: E: a	ACTE 906 mino	RIST ami aci	ICS: no a d	cids				• ,		i.		•
			(D)	TOP	OLOG	Y: 1	inea	r	* *.			•				0	8 *
	· (i	i) -M	OLEC	ULE	TYPE	: pr	oței	n	*				· 8	*	mar III		
	(x	i) S	EQUE	NCE	DESC	RIPT	ION:	SEQ	ID	NO:1	7:	. *			•	· · · · ·	•
Met	Val	Arg	Leu	Leu 5	Leu	Ile	Phe	Phe	Pro 10	Met	Ile	Phe ·	Leu	Glu 15	Met		
Ser	Ile	Leu	Pro 20	Arg	Met	Pro	Asp	Arg 25	Lys	Val	Leu	Leu	Ala 30	Gly	Ala	*	
Ser	Ser	Gln 35		Ser	Val	Ala	Arg 40	Met	Asp	Gly	Asp	Val 45	Ile.	Ile	Gly		
Ala	Leu 50		Ser	Val	His	His 55	Gln	Pro	Pro	Ala	Glu 60	Lys	Val	Pro	Glu	· ·	ŧ
Arg 65		Cys	Gly.	Gļu	Ile 70	Arg	Glu	Gln	Tyr	Gly 75	Ile	Gln	Arg	Val	Glu 80		*
	Met	Phe	His	Thr 85	Leu	Asp	Lys	Ile	Asn 90	Ala	Asp	Pro	Val	Leu 95	Leu	•	
Pro	Asn	Ile	Thr 100	Leu	Gly	Ser	Glu	Ile 105	Arg	Asp	Ser	Cys	Trp 110	His	Ser		
	•	115				Ser	120			•							
	130				. 1	133									Gly	• .	
Gln 145	Thr	Leu	Pro	Pro	Gly 150	Arg	Thr	Lys	Lys	Pro 155	11	Ala	Gly	Val	Ile 160		

## GURSTITUTE SHEET

		1000										•			
Gly-	Pro	Gly	Ser	Ser 165	Ser	Val	Ala	Ile	Gln 170	Val	Gln	Asn	Leu	Leu 175	Gln
Leu	Phe	Asp	Ile 180	Pro	Gln	Ile	Ala	Tyr 185	Ser	Ala	Thr	Ser	Ile 190	Asp	Leu
Ser	Asp	Lys 195	Thr	Leu	Tyr	Lys	Tyr 200	Phe	Leu	Arg	Val	Val 205	Pro	Ser	Asp
Thr	Leu 210	Gln	Ala	Arg	Ala	Met 215	Leu	Asp	Ile	Val	Lys 220	Arg	Tyr	Asn	Trp
Thr 225	Tyr	Val	Ser	Ala	Val 230	His	Thr	Glu	Gly	Asn 235	Tyr	Gly	Glu	Ser	Gly 240
Met	Asp	Ala	Phe	Lys 245	Glu	.Leu	Ala	Ala	Gln 250	Glu	Gly	Leu	Cys	Ile 255	Ala
HÌS	Ser	Asp	Lys 260	Ile	Tyr	Ser	Asn	Ala 265	Gly	Glu	Lys	Ser	Phe 270	Asp	Arg
Leu	Leu	Arg 275		Leu	Arg	Glu	Arg 280	Leu	Pro	Lys	Ala	Arg 285	Val	Val	Val
•	290				Met	295		8			300		. •		
Arg 305		Gly	Val	Val	Gly 310	Glu	Phe	Ser	Leu	Ile 315	Gly	Ser	Asp	Gly	1rp 320
				325			,		330	•	è				•
Gly	Ile	Thr	Ile 340	Lys	Leu	Gln	Ser	Pro 345	Glu	Val	Arg	Ser	Phe 350	Asp	Asp
Tyr	Phe	Lev 355		Lev	Arg	Leu	Asp 360	Thr	Asn	Thr	Arg	Asn 365	Pro	Trp	Phe
Pro	G10 370		Trp	Glr	His	Arg 375	Phe	Gln	Cys	Arg	380	Pro	Gly	His	Leu
Leu 385		a Asr	Pro	) Asr	Phe 390	Lys	Lys	Val	. Cys	395	Gly	) Asn	Glu	Ser	100
		÷		405	5			,	410	,			-	724	
	`-		420		•		-	425	<b>)</b>	• .		:	431	•	Cys
		43!	5				440	, ,							/ Arg
Lys	3 Let 450		ı Ası	Pho	e Leu	1 Ile 455	a Lys	s Sei	Se	r Phe	a Val	L G13	y Val	L Sei	Gly

465					<b>Asp</b> 470					•				•	
	•			485	Tyr				470				*		
		. <b>-</b>	500		Glu		•	100							
		515		•	Gly	100	520		٠.	, .	-				
	530	• •	•	٠	Lys	232					, -				
Trp 545	Ile	Cys	Thr	Ala	Cys 550	Lys	Glu	Asn	Glu	Phe 555	Val	Gln	Asp	Glu	Phe 560
Thr	Cys	Arg	Ala	Cys 565	Asp	Leu	Gly	Trp	Trp 570	Pro	Asn	Ala	Glu	Leu 575	Thr
Gly	Cys	Glu	Pro 580	Ile	Pro	Val	Arg	Tyr 585	Leu	Glu	Trp	ser	Asp 590	Ile	Glu
Ser	Ile	Ile 595	Ala	Ile	Ala	Phe	Ser 600	Cys	Leu	Gly	Ile	Leu 605	Val	Thr	Leu
Phe		Thr	Leu	Ile	Phe	Val 615	Leu	Tyr	Arg	Asp	Thr 620	Pro	Val	Val	Lys
Ser 625	Ser		Arg	Glu	Leu 630	Сув	Tyr	Ile	Ile	Leu 635	Ala	Gly	Ile	Phe	Leu 640
Gly	Tyr	Val	Cys	Pro 645	Phe	Thr	Leu	Ila	Ala 650	Lys	Pro	Thr	Thr	Thr 655	Ser
Сув	Tyr	Leu	Gln 660	Arg	Leu	Leu	Val	Gly 665	Leu	Ser	Ser	Ala	Met 670	Cys	Tyr
Sez	Ala	Lev 675	val	Thr	Lys	Thr	Asn 680	Arg	Ile	Ala	Arg	11e	Lev	Ala	Gly
Ser	Lys 690	Lys			Cys	Thr 695	Arg	Lys	Pro	Arg	Phe 700	Met	. Set	: Ala	Trp
A18	Gln		Ile	Ile	Ala 710	Ser	Ile	Leu	Ile	Ser 715	val	Glr	Leu	Th:	720
va]	Val	Thi	Lev	11e	e Ile	Met	: Glu	Pr	Pro 730	Met	: Pro	Ile	Let	735	Tyr
Pro	Ser	: Ile	Lys 740	. Glu		туг	: Lev	11e 745	Cys	a Asi	Thi	: S 1	750	i Lev	a Gly
Va:	L Val	75!	a Pr		l Gly	r Tyr	760	Gly	Lev	ı Let	ı Ile	76!	t Se	r Cyt	s Thr

									•								•	
	Tyn	Tyr 770	Ala	Phe	Lys	Thr	Arg 775	Asn .,	Val	Pro	Ala	Asn 780	Phe	Asn	Glu.	Ala		
	Lys 785	Tyr	Ile	Ala	Phe	Thr 790	Met	Tyr	Thr	Thr	Cys 795	Ile	Ile	Trp	Leu	Ala 800		
	Phe	Val	Pro	Ile	Tyr 805	Phe	Gly	Ser	Asn	Tyr 810	Lys	Ile	Ile	Thr	Thr 815	Cys	• •	
				820	,				825							Thr		
	Pro	Lys	Met 835	Tyr	Ile	Ile	Ile	Ala 840	Lys	Pro	Glu	Arg	* Asn 845	Val	Arg	Ser	*	
	Ala	Phe 850		Thr	Ser	Asp	Val 855	Val	Arg	Met	His	Val 860	Gly	Asp	Gly	Lys	er de	
	Leu 865		Cys	Arg	Ser	Asn 870	Thr	Phe	Leu	Asn	Ile 875	Phe	Arg	Arg	Lys	Lys 880		•
	Pro	Gly	Ala	Gly	Asn 885	Ala	Lys	Lys	Arg	Gln 890	Pro	Glu	Phe	Ser	Pro 895	Ser	•	
				Pro 900		• •			905	Leu		*						•
	(2)		.) SE ( (	QUEN A) L B) T C) S	CE C ENGI YPE:	HARA H: 4 nuc	CTER 095 :leic IESS:	ISTI base aci	CS: pai	.rs		*	•		:			
		(11	) MC	LECU	ILE I	YPE:	cDN	A				*		;				•
			c) FI	MEDI (B) ( EATUI (A) I (B) I	CLONI RE: NAME/	(KEY:	130 : CDS	5 33:	198	Î.	*		·.					•
	٠.	-									NO: 18				, •		, , , , , , , , , , , , , , , , , , ,	
٠																rccccc		60
																AGTGAC		.20
	GG	AGCCI	ATTC	TCT	CCTC	GTT (	GATA	agat	TC C	CTAC	CAGG	A TA	GGAG	CCTA	TCT	CCCTTT	2 1	.80
•																CTCCTG		40
		an an	01.00													TCTACT		300

	•									7 &	- ×			•	•		~	
					CC		አሮልጥ(	стст	GGA	ATGG(	GCC .	ACCC	ÇGTŢ(	GG C	CTGC	TGCAT		360
,	AAAC	ATTA	AA G	AGGA	GGAG.	3 GG.	CCTG	CCCA	GAT	CATG	GGA	CCCA	GGGC	CT G	CTAC	GGCTA	,	420
,	TACT	GAAC	CT G	CGCT	Gree	A CA	CGIG		••••				<b>-</b> -	<u> </u>	· · · · ·	NAC .		474
	GGAG	CGGG	GC C	CAGT	ATTC	A TG	GGTC	TCTA	GGC	CTTT	CCG ,	AA A' M	TG T et S	er G	ly I	rve		
						:				•••••			1					
				; -			mác		cee	CTG	CCC	CTC Leu	TGC.	CTA	CTC	CTC		522
	GGA	GGC	TGG	GCC	TGG	TGG.	Tro	Ala	Arg	Leu	Pro	Leu	Cys	Leu	Leu	Leu		
	GTÅ.	GIÀ	Trp	WIG.	115	10		1	· · · · · · ·		15	1	•		, .	20		
		***						. , 	መሮል	<b>ጥ</b> ሮር-	TTG	GGA	AAG	CCC	AAG	GGT Gly		570
	AGC	CTT	TAT	GCC	CCC	TGG	GIG Val	Pro	Ser	Ser	Leu	Gly	Lys	Pro	Lys	Gly		
	Ser	Leu	TAL	VTG	25	112	•			30		**			35			
		. ·.,	•					003	N COUCT	GAC :	ccc	GAC	ATC	ACA	CTG	GGA		618
	CAC	CCC	CAC	ATG	AAC	TCT	ATC	Ara	Ila	ASD	Gly	Asp	Ile	Thr	Leu	GGA Gly	•	
	His	TO	H1S	Met 40	ASI	261	1.10		45					50	•			0
	*:			••	•				خوما	mem	CAG	CCT	AAG	GCC	TGC	GGG		666
-	GGC	CTG	TTT	CCC	GTC	CAC	GGC	CGT	GGC	Ser	Glu	Gly	Lys	Ala	Cys	GGG Gly		
	Gly	Leu	Phe 55	Pro	ATT	UTR	GLY	60			÷		65	e* .				
	.*				· · ·			1 .		000	CTIC	GAG	GCC	ATG	CTG	TTT		714
	GAG	CTG	AAG	AAG	GAG	AAA	GGC	ATC	CAC	Ara	Leu	Glu	Ala	Met	Leu	TTT Phe		
	Glu	Leu	Lys	Lys	GIU	rys	75	116	1120	· • • • • • • • • • • • • • • • • • • •		80		• •				
		70	•	:		• :			·		Off 3	- CTIC	CCC	AAC	ATC	ACG		762
	GCC	CTG	GAÇ	CGC	ATC	AAC	AAT	GAC	. CCG	GAC Asn	Leu	Leu	Pro	Asn	Ile	ACG Thr 100		
	Ala	Leu	Asp	Arg	Ile	ASN 90	ASII	ASP	PLO	*****	95		*			100	٠,	
	85		7		,		. *						A CC	CAC	: GCC	CTG		810
	TTG	GGC	GCC	CGC	ATT	CTG	GAC	ACC	TGC	TCG	AGG	ASD	Thr	His	Ala	CTG Leu		-
	Leu	Gly	Ala	Arg	Ile 105	Dén	Asp	THE	Cyp	110					115	Leu	•	٠.
		• •		. 1								636	336	GAG		ACG		858
	GAG	CAG	TCA	CTG	ACC	TTT	GTG	CGG	GCG	CTC	ATC	Glu	LVS	Asp	Gl	C ACG		
	Glu	Gln	Ser	Leu	TUL	Phe	Val	Arg	125				, — <b>-</b>	130		y Thr		•
				120		,:	0	•					3.00	2 2 2 6	: CC	GAA		906
	GAG	GTC	CGC	TGC	GGC	AGG	CGG	GGC	CCG	CCC	ATC	: ATC	Thr	Lys	s Pr	C GAA	T w	
	Glu	Val	Arc	Cys	Gly	Arg	Arg	Г. G1У 140		, PLO			145	•	•	o Glu		00
			135	) [		•			•			•	•					954
	CGZ	GTG	GTO	GGI	GTC	ATI	GGA	GCT	TCG	GGG	AGC	TCC	· Val	Se	r Il	C ATG e Met		
	Arg	Val	. Val	L Gly	val	Il€	3 GT		Sei	GTA	. 261	160	)		¯·.	e Met		
		150	•													•		1002
	CTY	: 600	: AAC	C ATC	CTC	CGC	CTC	TTC	AAC	ATC	CCI	r CAG	ATC	Se	r TV	T GCC r Ala		
	Va.	Ala	Ası	ı Ile	Lev	r. wr.	,	ı- Phe	a Lys	3 TTE	175	5. 5.	7		1	r Ala 180		
	169	\$	•	į		1/	<b>J</b> .				•	•	•					1050
	, mc/	200	: 60	כ ככי	r Gac	TTC	G AGI	r GA	CAAC	AGG	CGC	CTAT	GAC	TT	e Ph	C TCC e Ser	2 -	
	Se	Thi	Al	a Pro	) AS	TIG!	u Sei	As	p Ası	sei 101	r arg	à TÀI	. nol		19	e Ser 5		
				,	TO	,												. 1098
			مرابق ت	6 CC	c TCI	A GA	C AC	A TA	C CA	G GC	CA	G GC	TA C	GT	G GA	TTA T	بد :	
	افات	e cil	a GT	~ ~~·										•	•			

## - PETITUTE SHEET

					;				, 1			_			_	-1-1	•	0
		•		200		Asp '			203					. *		. •		
	GTC	CGA	GCC	CTC	AAG	TGG	AÁC '	TAT	GTG	TCC	ACA.	CTG	GCC	TCA	GAG	GGC	· .	1146
	Val	Arg	Ala 215	Leu	Lys	TGG /	ASI	Tyr 220	Val	Ser .	Thr	Leu	225°	ser	GIU	GIY		•
		ma c	C 3M	CAG	አርጥ	GGT	GTG.	GAG	GCC	TTT	ATC	CAG	AAG	TCC	CGA	GAG		1194
	Ser	Tyr	Gly	Glu	Ser	GTA	Val 235	Glu	Ala	Phe	Ile	Gln 240	Lys	Ser	Arg	Glu		
		230		*				-1-6	mac	CITIC	AAG	ልጥጥ	CCA	cee	GAA	CCC		1242
	AAC	GGA	GGT	GTG Val	TGC	ATT Ile	Ala	Gln	Ser	Val	Lys	Ile	Pro	Arg	Glu	Pro	•	
	245	,			•	250	•				200		٠.		٠,			
	AAG	ACG	GGG	GAG	TTC	GAC	AAG	ATC	ATC	AAA	ÇGC	CTA	CTG	GAA	ACA	TCC Ser	•	1290
	Lys	Thr	Gly	Glu	Phe	Asp	Lys	Ile	Ile	Lys 270	Arg	Leu	Leu	GIU	275	Ser	· .	-
	<b>3</b> .		•	•	265							<b>410</b>	<b>63.6</b>	» mc	300	). NGG		1338
	AAT	GCC	AGG	GGT	ATC	ATC Ile	ATC	TTT	GCC Ala	AAC	GAG Glu	Asp	Asp	Ile	Arg	agg arg		:
	Asņ	ALA	Arg	280	TIE	116	116	- 110	285	-		_	·	290			•*	
			63.6	CCA	ССТ	CCC	AGG	GCC	AAC	CAG	ACC	ĠGC	CAC	TTC	TTT	TGG		1386
	GTG Val	Leu	Glu	Ala	Ala	Arg	Arg	Ald	You	Gln	Thr	Gly	His 305	Phe	Phe	Trp	•	
		¥.	295					200		•		•				•		1424
	ATG	GGT	TCI	GAT	AGC	TGG	GGC	TCC	AAG	AGT	GCC	CCT	GTG Val	CTG LAU	CGC	CTT		1434
	Met	Gly 310	Ser	Asp	Ser	Trp	Gly 315	Ser	гув	" Ser	WIG	320	102			Leu		
							663	cmo	. ACC	· Ariver	CTC	CCC	AAG	AGG	ATG	TCT	*	1482
	GAG	GAG	GTG Val	GCC Ala	GAG Glu	GGC	Ala	Val	Thr	Ile		Pro	Lys	Arg	Met	TCT Ser 340		
	325					3,30					,,,,		•			<u>.</u>	•	
	टनम	CGZ	GGG	TTC	GAC	CGA	TAC	TTC	: TCC	AGC	CGC	ACG	CTG	GAC	AAC	AAC Asn		1530
٠	Val	Arc	Gly	Phe	Asp	Arg	Tyr	Phe	Ser	Ser 350	*** 7	Thr	Leu	Asp	35	n Asn		ЭНС •
	0.	٠, ،			345			· <u>.</u>		.,		63.0			· car	r TGC		1578
	AGC	CG	AAC	ATC	TGG	TTT	GCC	GAG Glu	TTC	TGG Tro	GAG Glu	Asp	Asn	Phe	His	TGC Cys		
	Arc	Arg	J ASI	360	)	, PHG	, Ata		365	5	-	٠٠.	·	370	)			•
					CAC	Coc	стс	AAC	S AAC	G GGA	AGC	; ÇAC	: ATC	AAC	AA	G TGC	. *	1626
	LV	Le	G AG	r Arg	His	Ala	Leu	TA:	a ma	s Gly	Ser	His	: Il€ :385	Lys	Ly	s Cys		
			37!	5	•		•	300	,	•				;			<i>.</i> • •	1674
	AC	C AA	C CG	A GAC	G CGC	ATC	: GGG	CAC	G GA	C TCC	GCC	TAI	GAC	CAC	G GA	G GGG u Gly		16/4
	Th	r As	n Ar	g Glı	ı Arç	j Ile	395 395	GTI	n As	b ser	WIC	400				u Gly		
		39		-					m ~m	ሮ ጥል	<u>።</u> ።	- 2 ልጥር	G GG	C CA	C GC	G CTG	*	1722
	AA	G GT	G CA	G TT( n Ph	C GT( B Vai	3 ATT	GAC Ast	Al	a Va	1 Ty	r Ala	Met	E GI	y Hi	s Al	G CTG a Leu 420		•
	40	5		•		411	•			•		-					,	
	CA	C GC	C AT	G CA	C CG'	T GAG	CTC	TG	T CC	C GG	C CG	C GT	A GG	A CT	CIG	c cci	:	1770
	Hi	s Al	a Me	t Hi	s Ar	g Ası	p Lev	, Cy	9 FL	0 41	,,	<b>9</b> -	T GT	A THE	u cy	s Pro		•
										75	CHF	ET				•		•

								· .	•	7 14							•	
				, .	>	•			*	430	• •				435	• 0		
					425	٠.									•		*	
					GTG				C2C (	CTIC	نلىنلى	AAG	TAC	ATC	AGG	AAC	•	1818
	CGC	ATG	GAC	CCC	GTG Val	GAT (	GGC (	ACC	CAG		TAII	Lvs	Tyr	Ile	Arg	Asn	,	
	Arg	Met	Asp	Pro	Val	Asp (	Gly '	THE		reu	Dea		-3-	450	-		,	
,				440					• • •	. 10				. *				
					GGC	: *				COM	CTA	ACC	TTC	AAT	GAG	AAC	-	1866
	GTC	AAC	TTC	TCA	GGC	ATT.	GCG	GGG	AAC	CCI	V= 1	Thr	Phe	Asn	Glu	Asn "	,	
•	Val	Asn	Plie	Ser	Gly	Ile .	VTG	<b>-</b> 1	ASN	PEU	VAI		465		* 1	•		
			455					700									. ,	ė
	-	•			GGG	. 1				M3.0	CNG	TAC	CAA	CTG	CGC	AAT	•	1914
	GGA	GAC	GCA	CCG	GGG	CGÇ	TAC	GAC	ATC	TAC	Cla	Tyr	Gln	Leu	Arq	Asn	•	
	Glv	ASD	Ala	Pro	Gly	Arg	Tyr	Asp	TTG	TAT	GIM	480			•	Asn		
; ;		470	•			. •	4/3				- '					ē		
		• • •			TAC					mac .	mcc.	ACA	GAC	CAC	CTG	CAC .		1962
	GGC	TCG	GCC	GAG	TAC	AAG	GTC	ATC	GGC	TCG	TGG	Thr	ASD	His	Leu	His		
	Glv	Ser	Ala	Glu	TAC	Lys	Val	Ile	GTÅ	Ser	TEP	.1111	nop.			500		•
	485			•		490	-			9 .	493							
	400		•	•							3 CM	ccc	CAG	CAG	CTG	CCG		2010
	CTC	AGA	ATA	GAG	CGG	ATG	CAG	TGG	CCA	GGG	AGT	Clv	Gln	Gln	Leu	Pro		
	TAII	Arg	Ile	Glu	CGG	Met	Gln	Trp	Pro	2	ser	GTA	G1	·	515			
					505		ı		•	510	F	· · · · · · · · · · · · · · · · · · ·						•
						•					000	CAG	CGA	AAG	AAG	ACT		2058
	CGC	TCC	ATC	TGC	AGT	CTG	CCC	TGC	CAG	CCC		Clu	Ara	T.VS	Lvs	Thr	•	
	7-4	Ser	Tle	CVS	AGT Ser	Leu	Pro	Cys	~	Pro	GIY	GIU	ur A	530		٠		
	Arg			520		-	,	0	525	~		•					· · · .	• •
	•											-	TGC	ACC	GGG	TAC		2106
	CMC	AAG	GGC	ATG	GCT	TGC	TGC	TGG	CAC	TGC	GAG		Tac	Thr	Glv	TVY	. 0	.,
•	77-1	Tve	Glv	Met	GCT	Cys	Cys	Trp	His	Cys	GIU	LED	CYS	****		2 -		
	Val	ny s	535		,	*		540			-		545	•	re.	8-	٠.	• 41
	•			•			``;		_					TAC	GAC	ATG		2154
	G3.C	TO C	CAB	GTG	GAC	CGC	TAC	ACC	TGT	AAG	ACC	TGC	7-0	TAC	) Aen	ATG Met		÷
	CAG	TVC	Gla	Val	Aso	Arq	Tyr	Thr	Cys	Lys	Thr	Cys	PLO	TÄT	, asp	Met		•
	GTU	550	GII				555	,		•		560		-			•	
		220		•		•		•	٠					NTC.	• СТС	AAG	141	2202
1,	000	d00	3.03	GAG	: AAC	CGC	ACG	AGC	TGC	CAG	CCC	ATC	7	TIC	, U=1	AAG Lys		
	- CGG	, CCC	The	Gli	ASD	Arg	Thr	Ser	Cys	Gln	Pro	TTE	Pro	. T.T.	, ,,,,	Lys 580		•
٠.			, 1111	. 41		570			1. :		575			* *		. 500	•.	•
	565				•	- •								~~		י כתה		2250
			. mcc	CAC	- ጥርር	CCG	TGG	GCC	GTG	CTG	CCC	CIC	TTC	CTC	. 31.	GTG Val		× .
•	TTG	GAG	100	) GAN	Ser	Pro	Trp	Ala	Val	Leu	Pro	Let	l Phe	Lei	E WIG	val	•	7
•	Leu	f GT	i iti	, ASI	ب. 585 8. پر			٠.		- 590	··	,			27:	•		
;			•					• • •	••		0			- cm/		TAC		2298
T			- 2 <i>m</i> v	- GC(	- 600	ACG	CTG	TTC	GTG	GTG	GTC	ACC	11.1	GI	3 CG	TAC		
•	GIG	GGG	, AI	S Blo	Ala	Thr	Leu	Phe	Val	. Val	. Val	LThi	r Pn∈	2 va.	r wr	Tyr		
	. Val	L GI	TŤ	60	n ni				605	•	-		•	OT.	U			
٠,	•								••		٠,			. 3.04	- ma	c cric		2346
					. አጥር	GTC	AAG	GCC	TCG	GGC	CGG	GA(	CIG	AG	CATA	c GTG r Val		
,	AAC	GA.	. AC	- D-	o TIA	Val	Lvs	Ala	Ser	Gly	Arc	g Gl	ı Let	i se	E TY	r Val		
	Ası	l Asi	ını	, rl'	0 114	- <del></del>		620	)			•	625	•	•		, -	
	<u>.</u> .		61	<b></b>	· ·						•					a 1ma	•	2394
					a = m/	- WW	י בידור	TG	TAC	GCC	C. AC	r ac	CIT	CI	C AT	G ATC		<i>と</i> J フ マ
	CTC	G CI	3 GC	G GG	CAT	2 DP-	TAI	CV	TVI	Ala	a Thi	r Th	r Pho	e Le	u Me	t Ile		•
	Let	ı Lei	ı Al	a Gl	Å TΤ.	s Pile	, ,,,,,,	5 ···	,-		•	.64	0		•			
		671	n .														•	2442
							2 200	- 44 <u>5</u> 4	r TC	CT	C CG	C CG	C AT	C TT	CCI	A GGG		4444
•	GC	A GA	G CC	g ga	C CI	. GGC	, mb.	p	- CA1	LA	u Ar	g Ar	g Ile	e Ph	e Le	A GGG		
	Al	a Gl	u Pr	o As	p Le	r Gri		L Cy	. JE		65	<b>5</b>				u Gly 660		-
	64	5			÷	650	<b>,</b>			•								7
	-									4								-

	GCC	: AAA	TCA	GAG	CIG	TGT	GAU	AAC	10	, uno		· .		SUE	STIT	UTE	SHE	ET	
		870					879 GAG	• • •	• (•T•)	: האה	. ACC	- ×		CTC	GC:	r acc	. *	3162	
	ATG Met	Ser	Asn	AAG Lys	TTC Ph	ACA Thr	Gl	, råa	GGC	AAC Asn	Phe	AGG Arg 880	PLU	AAT Asn	GGG	GAA Glu		3114	•
•	Val	Pro	Lys 855	Arg	Lys	Arg	Ser	860	rys	. WTG	, ,val	VQI	865						
	GTG.	CCC	AAG			CGC	AGT	CIC		. GCC	GTG	GTC	ACC	GCC	GCC	ACC Thr		3066	
	CTC	TAC	ATG Met	CCC Pro 840	Lys	Val	Tyr	Ile	Ile 845	Leu	Phe	His	Pro	Glu 850	GI	AAC Asn	*	<del></del>	0
	*	•			825					830	•		•			•		3018	
•	ACA	CTG Leu	ACG Thr	GTÇ Val	Ser	Val	AGT	CTG Leu	AGC Ser	, VIS	TCA Ser	GTG Val	TCC	CTG	GGG Gly 835	ATG Met		2970	
	805		*		•	810		· · ·			gTo	•				02,0		2070	
	TTT	TTT	GGC	ACC	TCA	CAG	TCA	GCC	GAC	AAG	CTG	TAC	ATC Ile	CAG Gln	ACA Thr	ACC Thr		2922	
	Phe	Thr 790	Met	Tyr	Thr	Thr	Cys 795	Ile	Val	Trp	Leu	Ala 800	Phe	Ile	Pro	Ile			
:	<b>ም</b> ሞር	ACC	775 ATG	MAC	ACC	ACC	TGC	יויייע	GTC	TGG	CTG	GCC	TTC	ATC	CCC	ATC		2874	
	AAG Lys	ACC Thr	Arg	Gly	GTG Val	CCC	GAG Glu	ACC Thr 780	TTC	AAC	GAG	Ala	Lys 785	Pro	Ile	GGC	*	2020	,
				760	,				765	-	•			,,,				2826	•
••	CTG	GGC	TAC	AGC	ATG Met	CTG	CTG	ATG Met	GTC Val	ACG	TGT Cys	ACT	GTG Val	TAC	GCC Ala	ATC	ing or a g	2778	
	Arg	Gly	Val	Leu	Lys 745	Cys	Asp	Ile	Ser	<b>Asp</b> 750	Leu	Ser	Leu	116	755	ren	<i>.</i>	 	
	3.00	GGC	GTG	CTC	AAG	TGC	GAC	ATC	TCG	GAC	CIG	TCC	CTC	ATC	TGC	CTG		2730	:
	TCG Ser 725	GTG al	GTG Val	GAC Asp	TTC Phe	CAG Gln 730	GAC Asp	CAA Gln	CGG	ACA	Leu 735	Asp	Pro	Arg	Phe	Ala 740	•	2002	ć
		710		•			715	•		•		/20	ě		1		*	2682	,
	CTG	CAG	CTG	CTC	GGC	ATC Ile	TGC CVs	GTG Val	TGG	TTC Phe	GTG Val	GTG Val	GAC Asp	CCC Pro	TCC Ser	CAC	* .	2634	
	Phe	Ile	Ser 695	Pro	Ala	Ser	Gln	Leu 700	Ala	Ile	Thr	Phe	11e 705	Leu	Ile	Ser			
	ттс	ATC	AGC	680. CCG	CCC	TCG	CAG	CTG	GCC	ATC	ACC	TTC	ATC	CTC	ATC	TCC	e	2586	
	ATT Ile	TAC Tyr	CGC Arg	Ile	Phe	GAG Glu	CAG Gln	GGC Gly	AAA Lys 685	CGG	TCG Ser	GTC Val	AGT	GCC Ala 690	Pro	Arg		2538	
	•				665	·		•		6/0					0/5		*		
,	CIC	GGC	ATG	AGC	ATC	AGC Ser	TAC	GCG.	GCC	CTG	CTG Len	ACC Thr	AAG Lvs	ACC	AAC Asn	CGC		2490	
								•	· .	•					, .				

Ala-Lys Ser Glu Leu Cys Glu Asn Leu Glu Thr Pro Ala Leu Ala Thr 895 900	
AAA CAG ACC TAC GTC ACC TAC ACC AAC CAT GCC ATC TAGCCGGGCC Lys Gln Thr Tyr Val Thr Tyr Thr Asn His Ala Ile 905	3208
GCGGAGCCAA-GCAGGCTAAG GAGCCACAAC CTCTGAGGAT GGCACATTGG GCCAGGGCCG	3268
TTCCCGAGGG CCCTGCCGAT GTCTGCCCGC CTCCCGGGCA TCCACGAATG TGGCTTGGTG	3328
CTGAGGACAG TAGAGACCCC GGCCATCACT GCTGGGCAAG CCGTGGTGGG CAACCAGAGG	3388
AGGCCGAGTG GCTGGGGCAG TTCCAGGTTA TGCCACACAC AGGTCTTCCT TCTGGACCAC	3448
TGTTGGCCCA GCCCCAAAGC ACAGGGGCTC GGTCTCCAGA GCCCAGCCCT GGCTTCCTCT	3508
CCTTC CTCCT GCCTCCGTCT GTCCTGTGGG TGACCCCGGT TGGTCCCTGC CCCGTCTTTA	3568
COTTCTCTT CCGTCTTTGC TCTGCATGTG TTGTCTGTTT GGGCCCTCTG CTTCCATATT	3628
TITCATTCT GCTCCTGGCC TTCCCCTGCC ATCTGCCCTG CCCCCTGCCC CTCCTCCCTG	3688
AGCTGCCCCA TCCCCGCCAT CATTTTCTCT TCTGTTCCCC CTCGATCTCA TTTCCTACCA	3748
AGCTGCCCCA TCCCCGCCAT CATTITICION TO TOTAL ACCORDANA AGAGAAAAAA GCCTTCCCCC TACTTGGCTT CATCCACCAA CTCTTTCACC ACGTTGCAAA AGAGAAAAAA	3808
GCCTTCCCCC TACTTGGCTT CATCLACCAA CICTIONIO AAAAACTAAT CTTGAGTGTG	3868
AAAGGGGGG GGGAATCACC CCCTACAAAA AAGCCCAAAC AAAAACTAAT CTTGAGTGTG	3928
TTTCGAAGTG CTGCGTCCTC CTGGTGGCCT GTGTGTCCCT GTGCCTGCAG CCTGCTGCC	3988
CGCCCTACCC GTCTGCCGTG TGTCCTGCCC CCCCCGCCTG CCCGCCTTGC CCTTCCTGCT	4048
AACGACACGG AGTTCAGTGC CTGGGTGTTT GGTGATGGTC TCTGATGTGT AGCATGTCTG	409
TTTTTATACC GAGAACATTT CTAATAAAGA TAAACACATG GTTTTGC	

### (2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 912 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (XI) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met Ser Gly Lys Gly Gly Trp Ala Trp Trp Trp Ala Arg Leu Pr Leu
1 10 15

Cys Leu Leu Ser Leu Tyr Ala Pr Trp Val Pro S r Ser Leu Gly
25 30

Lys Pro Lys Gly His Pr His Met Asn Ser Ile Arg Ile Asp Gly Asp 40

SUBSTITUTE SHEET

		•			Tou	Dha	Pro	ا ا عال	His	Glv	Ara	Glv	Ser	Glu	Gly
	50			•	f t	55									
65		•	•		70				9	1.3				Leu	
				85					90				•	Leu 95	
Pro	Asn	Ile	Thr 100	Leu	Gly	Ala	Arg	Ile 105	Leu	Asp	Thr	Cys	Ser 110	Arg	Asp
Thr		Ala 115	Leu	Glu	Gln	Ser	Leu 120	Thr	Phe	Val	Arg	Ala 125	Leu	Ile	Glu
Ĺys	Asp 130		Thr	Glu	Val	Arg 135	Cys	Gly	Arg	Arg	Gly 140	Pro	Pro	Ile	Ile
Thr 145		Pro	G1u	Arg	Val 150	Val	Gly	Va1	Ile	Gly 155	Ala	Ser	Gly	Ser	Ser 160
Val	Ser	Ile	Met	Val 165	Ala	Asn	Ile	Leu	Arg 170	Leu	Phe	Lys	Ile	Pro 175	Gln
		٠.	180					182	41				190	•	
		195			r.		200	•				205		Gln	. :
	210					215	•				220		• •	Thr	
225	<b>;</b>		1	1	230	· ·	•	4		433	٠ .			Ile	
	- •			245			•		250						Ile
	•		260	<b>,</b>	•			265	•		•		2/0		Leu
		275					280					20.	;		Asp
	290	,		•		295	•	•		•. ·	300				Gly
. 30!	<b>5</b> .		į		310	)	:	4		31:	,				320
Va.	l Lev	ı Arç	, Lev	Glu 325	Glu	· Val	Ala	Glu	Gl ₃	y Ala	va]	Thi	rIle	Lev 335	Pr
Ly	s Arg	y Met	: Ser 340	val	L Arg	Gly	Phe	Ast 345	Arq	TY1	r Phe	s Set	r Sei 350	r Arg	Thr

				•			8									1	
		•	355	, *	•	Arg		360,			•		•	٠.			
`		370		•		Leu	3/5				÷			•			
	385		•			Asn 390					334	E 100		. •	•	•	
					405	Val		٠,		410	·* .	•		•	4		
				420		Ala	-		425							•	-
	. 4		435	٠, .	•	Met		440							•		
	Tyr	Ile 450	Arg	Asn	Val	Asn	Phe 455	Ser	Gly	Ile	Ala	Gly 460	Asn	Pro	Val	Thr	
	Phe	Asn	Glu	Asn	Gly	Asp 470	Ala	Pro	Gly	Arg	Tyr 475	Asp	Ile	Tyr	Gln	Tyr 480	•
	Gln	Leu	Arg	Asn	Gly 485	Ser	Ala	Glu	Tyr	Lys 490	Val	Ile	Gly	Ser	Trp 495	Thr	
	Asp	His	Leu	His 500	Leu	Arg	Ile	Glu	Arg 505	Met	Gln	Trp	Pro	Gly 510	Ser	Gly	
•	Gln	Gln	Leu 515	Pro	Arg	Ser	Ile	Cys 520	Ser	Leu	Pro	Cys	Gln 525	Pro	Gly	Glu	
.*	Arg	Lys 530		Thr	Val	Lys	Gly 535	Met	Ala	Cys	Cys	Trp 540	His	Cys	Glu	Pro	
	Cys 545		Gly	Tyr	Gln	Tyr 550	Gln	Val	Asp	Arg	Tyr 555	Thr	Cys	Lys	Thr	Cys 560	
	Pro	Tyr	Asp	Met 2	Arg . 565	Pro	Thr	Glu	Asn	Arg 570	Thr	Ser	Cys	Gln	Pro 575	Ile	
	Pro	Ile	Val	Lys 580	Leu	Glu	Trp	Asp	Ser 585	Pro	Trp	Ala	Val	Leu 590	Pro	Leu	
	Phe	Leu	Ala 595	Val	Val	Gly	Ile	Ala 600	Ala	Thr	Leu	Phe	Val 605	Val	Val	Thr	
	Phe	Val 610		Tyr	Asn	Asp	Thr 615	Pro	I1	Val	Lys	Ala 620	Ser	Gly	Arg	Glu	
	Leu 625	SI		Val	Lev	Leu 630	Ala	Gly	Ile	Phe	Leu 635	Cys	Туг	Ala	Thr	Thr 640	
			Met	Ile	Ala 645	Glu S	Pro	) Asp	Leu	Gly 650	Thr	Cys	Sei	: Lev	Arg 655	Arg	

# SUBSTITUTE SHEET

Ile	Phe	Leu	Gly 660	Leu	Gly	Met	Ser	Ile 665	Ser	Tyr	Ala	Ala	Leu 670	Leu	Thr
		675	1		Tyr		990				• '		••		
	690	•			Ile	673			·. ·						
705					Gln 710	•		•			•				•
			,	725	Val	. ,			/30	• '					
			740		Gly		•	745			•				
		755					760						•		Thr
٠.	770					//5				<i>:</i> .		•		_*	Ala
785			-		/90	• •	•	•		.,,,,	٠.	• •	•		Ala 800
	•	•		805	)·.	•		-	0.10	,					Tyr
			820				·	043		•	•	:			Val
		83:	•				040								His
	850	) '-			;	805		• :		•		. *	87		Val
86	5			<b>y</b> .	870	) .	-			• • • • • • • • • • • • • • • • • • • •	•		~	•	880
		•	. '	88	•			٠,	. 05			-	•		r Pro
Al	a Let	u Al	a Th	r Ly: O	s Gli	n Thi	r .Tyi	90!	1 .Th: 5	r Ty:	r Thi	r Asi	910	s Ala	lle

### (2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:

  (A) LENGTH: 2426 base pairs

  (B) TYPE: nucleic acid

  (C) STRANDEDNESS: single

  (D) TOPOLOGY: linear

### _ (ii) MOLECULE TYPE: cDNA

## (vii) IMMEDIATE SOURCE: (B) CLONE: SR13

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:	
CCCAACATCA CGTTGGGCGC CCGCATTCTG GACACCTGCT CGAGGGACAC CCACGCCCTG	60
GAGCAGTCAC TGACCTTTGT GCGGGCGCTC ATCGAGAAGG ACGGCACGGA GGTCCGCTGC	120
GGCAGGCGGG GCCCGCCCAT CATCACCAAG CCCGAACGAG TGGTGGGTGT CATTGGAGCT	180
TCGGGGAGCT CCGTCTCGAT CATGGTGGCC AACATCCTCC GCCTCTTCAA GATCCCTCAG	240
ATCACCTATG CCTCCACGGC CCCTGACTTG AGTGACAACA GCCGCTATGA CTTCTTCTCC	300
CGGGTGGTGC CCTCAGACAC ATACCAGGCC CAGGCCATGG TGGATATTGT CCGAGCCCTC	360
AAGTGGAACT ATGTGTCCAC ACTGGCCTCA GAGGGCAGCT ACGGTGAGAG TGGTGTGGAG	420
GCCTTTATCC AGAAGTCCCG AGAGAACGGA GGTGTGTGCA TTGCCCAGTC GGTGAAGATT	480
CCACGGGAAC CCAAGACGGG GGAGTTCGAC AAGATCATCA AACGCCTACT GGAAACATCC	540
AATGCCAGGG GTATCATCAT CTTTGCCAAC GAGGATGACA TCAGGAGGGT GTTGGAGGCA	600
GCTCGCAGGG CCAACCAGAC CGGCCACTTC TTTTGGATGG GTTCTGATAG CTGGGGCTCC	660
AAGAGTGCCC CTGTGCTGCG CCTTGAGGAG GTGGCCGAGG GCGCAGTCAC CATTCTCCCC	720
AAGAGGATGT CTGTTCGAGG GTTCGACCGA TACTTCTCCA GCCGCACGCT GGACAACAAC	780
AGGCGCAACA TCTGGTTTGC CGAGTTCTGG GAGGACAACT TCCATTGCAA GTTGAGCCGC	3 840
CACGCGCTCA AGAAGGGAAG CCACATCAAG AAGTGCACCA ACCGAGAGCG CATCGGGCAC	900
GACTCGGCCT ATGAGCAGGA GGGGAAGGTG CAGTTCGTGA TTGACGCTGT GTACGCCATC	960
GGCCACGCGC TGCACGCCAT GCACCGTGAC CTGTGTCCCG GCCGCGTAGG ACTCTGCCC	1020
CGCATGGACC CCGTGGATGG CACCCAGCTG CTTAAGTACA TCAGGAACGT CAACTTCTC	1080
GGCATTGCGG GGAACCCTGT AACCTTCAAT GAGAACGGAG ACGCACCGGG GCGCTACGA	1140
ATCTACCAGT ACCAACTGCG CAATGGCTCG GCCGAGTACA AGGTCATCGG CTCGTGGAC	A 1200
GACCACCTGC ACCTCAGAAT AGAGCGGATG CAGTGGCCAG GGAGTGGCCA GCAGCTGCC	G 1260
CGCTCCATCT GCAGTCTGCC CTGCCAGCCC GGGGAGCGAA AGAAGACTGT GAAGGGCAT	G 1320
GCTTGCTGCT GGCACTGCGA GCCCTGCACC GGGTACCAGT ACCAAGTGGA CCGCTACAC	C 1380
TGTAAGACCT GCCCCTACGA CATGCGGCCC ACAGAGAACC GCACGAGCTG CCAGCCCAT	C 1440
9.89: Branchis	/

• •				• •		
CCCATCGTCA	AGTTGGAGTG	GGACTCGCCG	TGGGCCGTGC	TGCCCCTCTT	CCTGGCCGTG	1500
GTGGGCATCG	CCGCCACGCT	GTTCGTGGTG	GTCACGTTTG	TGCGCTACAA	CGATACCCCC	1560
ATCGTCAAGG	CCTCGGGCCG	GGAGCTGAGC	TACGTGCTGC	TGGCGGGCAT	CTTTCTGTGC	1620
TACGCCACTA	CCTTCCTCAT	GATCGCAGAG	CCGGACCTGG	GGACCTGTTC	GCTCCGCCGC	1680
			TACGCGGCCC			1740
ATTTACCGCA	TCTTTGAGCA	GGGCAAACGG	TCGGTCAGTG	CCCCCCGTTT	CATCAGCCCG	1800
CCTCCCACC	TGGCCATCAC	CTTCATCCTC	ATCTCCCTGC	AGCTGCTCGG	CATCTGCGTG	1860
			GTGGACTTCC			1920
			GACATCTCGG			1980
					GACCCGAGGC	2040
			: ATCGGCTTCA			2100
			GGCACCTCAC			2160
ATTCAGACA	CCACACTGAC	GGTCTCCGT	AGTCTGAGCG	CTTCAGTGTC	CCTGGGGATG	2220
ATCCAGAGE	CCALAGTET	CATCATCCT	TTCCATATTI	TTCCATTCT	CTCCTGGCCT	2280
					CCCCCCCATC	2340
TCCCCTGCC	a respondence	теретельной	r TTCCTACCAG	COTTOCCO	ACTTGGCTTC	2400
0		*				2426
CTCCACCAA	C TCTTTCACC	A CGTTGC				, ,

#### (2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Arg Asp Ser Leu Ile Ser Ile Arg Asp Glu Lys Asp Gly Leu Asn Arg

Cys

### (2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

	*	(B)	TYPE:	amino 8	sindre	1					
, ,	(ii)	MOLEC	ULE TY	PE: pe	otide						
		•								*	
- 10	(xi)	SEQUI	ENCE DE	SCRIPT	ION: SE	O ID NO	:22:		T.em 1	. 1 - 1 A	ret V
	Asp 1	Arg	Leu Lev	Arg L	ys Leu	Arg Glu	Arg I	eu Pi	o riva :	1	5
2)	: Info	RMATI(	ON FOR	SEQ ID	NO:23:	*					
	(i)	(A) (B)	LENGTI TYPE: STRAN	H: 16 a amino	: singi	Tas					
			Y _i	YPE: pe							
	(xi) Glu 1	SEQU 1 Glu	ence D	ESCRIPI p Phe 1 5	'ION: Si Asp Glu	Tha egg	/ Asp 10	Ala P	ro Gly	Arg	Tyr A 15
(2)	INFO	ORMATI	ON FOR	SEQ II	NO:24			*			
	<b>(1</b> )	(A) (B)	LENGT TYPE:	H: 15 d	51 Slng						
			.m		eptide						
	(xi	) SEQ	uence i	DESCRIP	TION: S	EQ ID N	0:24:	<b>03.</b> 0 (	ili kan	Glu	Leu
· ·	Gl 1	u Phe	Val T	yr Glu 5	Arg Glu	Gly As	n Thr 10	GIU G	ita wah	914	15
(2)	INF	ORMAT	ION FO	R SEQ I	D NO:25	i <b>:</b> ,	**************************************				s - ()
		SEQ (A (B	UENCE LENG TYPE	CHARACT	ERISTIC amino a acid S: sing	:S: cids		· · · · · · · · · · · · · · · · · · ·			

SUBSTITUTE SHEET

'Val

103

- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Pro Glu Arg Lys Cys Cys Glu Ile Arg Glu Gln Tyr Gly Ile Gln Arg
1 5 10 15

- (2) INFORMATION FOR SEQ ID NO:26:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 16 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: peptide
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Ile Gly Pro Gly Ser Ser Ser Val Ala Ile Gln Val Gln Asn Leu Leu
10 15

- (2) INFORMATION FOR SEQ ID NO:27:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 15 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: peptide
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Ile Ala Tyr Ser Ala Thr Ser Ile Asp Leu Ser Asp Lys Thr Leu
1 10 15

- (2) INFORMATION FOR SEQ ID NO:28:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 16 amino acids
    - (B) TYPE: amin acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: peptide

(2)

104

			Y											.* •	
	(xi) S	SEQUE	NCE DI	SCRII	PTION	: SEQ	ID N	10:2	8:	* :			*		
	Lys I	Lys P	ro Gly	y Ala 5	Gly	Asn A	la Ly	s I	ro 'Aa	Arg	Gln	Pro	Glu	Phe 15	
:			•	,			*		1,				•	• .	
2)	INFOPI						0			100					
	(i) :	(A) (B)	NCE CLENGT LENGT TYPE: STRAN	H: 16 amin DEDNE	amın o aci ss: s	d ingle	us	*			, ,			- 1	•
•		•	TOPOL				•	٠.,.							
	(ii)	MOLEC	ULE T	YPE:	pepti	.de	-				• .				
			٠.	• .	,				· ·	'	•				,
	(xi),	SEQUI	ence d	ESCRI	PTION	: SEC	] ID	NO:	29:	-					
, .	Pro	Glu !	Phe Se	r Pro	Ser	Ser (	31n C	ys	Pro 10	Ser	Ala	His	Ala	. Glr 15	ľ
	1	\$ °		<b>5</b> .					¥.	•		*		Ε΄.	
(2)	INFOR	ITAM	ON FOR	SEQ	ID NO	0:30:	* **			• 0 •	1,0		a see	``	
	<b>(1)</b>	/ <b>A </b>	ENCE C	TH: 1:	) ami	UO ar	: ids		9.9			9			٠.
٠, ٠		. (C)	TYPE: STRAI TOPO	NDEDNI	ESS: - :	sindr.	e	ı			*	÷	TO .		
	(ii)	MOLE	CULE :	TYPE:	pept	ide			* ,	•	*.	0			,
				-			*					*	• • •		
	(xi)	SEQU	ENCE	DESCR	IPTIO	n: se	Q ID	NO	30				*	Ť	• ;- ;:
•	Asp 1	Lys	Ile I	le Ly 5	s Arg	Leu	Leu (	Glu	Thi 10	r Se	r As	n Al	a Ar	g Gl 15	y
(2	) INFO	RMATI	ON FO	R SEQ	ID N	0:31:	;	· .		•	•				
	(主)	(A) (B)	JENCE LENG TYPE STRA	TH: 1 : ami NDEDN	o ami no ac ESS:	no ac id singl	iTris							· :	
٠,	0	(a)	TOPO	LOGY:	line	ear		:							.,

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Val Asn Phe Ser Gly Ile Ala Gly Asn Pro Val Thr Phe Asn Glu Asn 1 5

- (2) INFORMATION FOR SEQ ID NO:32:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 16 amino acids
      - (B) TYPE: amino acid
      - (C) STRANDEDNESS: single
      - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: peptide
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Gly Glu Ala Lys Ser Glu Leu Cys Glu Asn Leu Glu Thr Pro Ala Leu 1 10 15

- (2) INFORMATION FOR SEQ ID NO:33:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 17 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: peptide
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Pro Ala Arg Leu Ala Leu Pro Ala Asn Asp Thr Glu Phe Ser Ala Trp

Val

15

20

25

30

#### WHAT IS CLAIMED IS:

- 1. An isolated mammalian G protein-coupled glutamate receptor or a fragment thereof.
- 2. The G protein-coupled glutamate receptor of claim
  1, which is substantially pure.
- 3. The G protein-coupled glutamate receptor of claim
  1, which is human or rodent.
  - 4. An antiserum obtained from an animal immunized with the G protein-coupled glutamate receptor of claim 1.
- 5. A monoclonal antibody which specifically binds to the G protein-coupled glutamate receptor of claim 1.
  - 6. The G protein-coupled glutamate receptor of claim 1, which binds glutamate or quisqualate and thereby activates phospholipase C or stimulates inositol phospholipid metabolism in a vertebrate cell.
  - 7. A recombinantly produced polypeptide having the activity of a mammalian G protein-coupled glutamate receptor.
  - 8. The polypeptide of claim 7, which has the activity of a human or rodent mammalian G protein-coupled glutamate receptor.
- 9. An isolated and purified polynucleotide molecule which codes for a mammalian G protein-coupled glutamate receptor or a fragment thereof.
- 10. The polynucleotide of claim 9, which is a genomic DNA sequence, a cDNA sequence, or an RNA antisense sequence.

- 11. The polynucle tide of claim 9, which codes for human or rodent G protein-coupled glutamate receptor.
- 12. The polynucleotide of claim 9, which encodes a polypeptide displaying mammalian G protein-coupled glutamate receptor activity.
- 13. The polynucleotide of claim 9, which is substantially the sequence of Fig. 5, Fig. 7, Fig. 8 or Fig. 9.

14. A probe which comprises an oligonucleotide capable of specifically hybridizing with a gene which encodes a mammalian G protein-coupled glutamate receptor or a fragment thereof.

15

- 15. The probe of claim 14, which comprises from about 40 to about 60 nucleotides in length.
- 16. The probe of claim 15, which is labeled to provide a detectable signal.
  - 17. A DNA construct comprising the following operably linked elements:
    - a transcriptional promoter;

25

- a DNA sequence encoding a mammalian G protein-coupled glutamate receptor or a fragment thereof; and
  - a transcriptional terminator.
- 18. The DNA construct of claim 17, wherein the DNA sequence encodes a human or rodent G protein-coupled glutamate receptor polypeptide.
  - 19. The DNA construct of claim 17, wherein the DNA sequence encoding the mammalian G protein-coupled glutamate receptor is substantially the sequence of Fig. 5, Fig. 7, Fig. 8 or Fig. 9.

10

15

20

30

- 20. A cultured eukaryotic cell transformed or transfected with a DNA construct which comprises the following operably linked elements:
  - a transcriptional promoter;
- a DNA sequence encoding a mammalian G protein-coupled glutamate receptor or a fragment thereof; and a transcriptional terminator.
- 21. The eukaryotic cell of claim 20, which is a mammalian cell.
- 22. The eukaryotic cell of claim 20, which does not express endogenous G protein-coupled glutamate receptors.
- 23. The eukaryotic cell line of claim 20, wherein the DNA sequence encodes a human or rodent G protein-coupled glutamate receptor polypeptide.
- 24. The eukaryotic cell line of claim 21, wherein the G protein-coupled glutamate receptor polypeptide encoded by the DNA sequence is coupled to G protein in a mammalian cell.
  - 25. The DNA construct of claim 20, wherein the DNA sequence encoding the mammalian G protein-coupled glutamate receptor is substantially the sequence of Fig. 5, Fig. 7, Fig. 8 or Fig. 9.
  - 26. A method for producing a mammalian G proteincoupled glutamate receptor, which comprises:
- growing eukaryotic cells transformed or transfected with a DNA construct which comprises a DNA sequence coding for the expression of the G protein-coupled glutamate receptor, and isolating the receptor from the cells.
- 27. The method of claim 26, wherein the cells are cultured mammalian cells.

WO 92/10583

- 28. Th method of claim 26, wherein the glutamate receptor is human or rodent.
- 29. The method of claim 26, wherein the glutamate receptor is isolated by immunoaffinity purification.
- 30. The method of claim 26, wherein the G protein-coupled glutamate receptor is not coupled to protein G in the eukaryotic cells.

10

15

5

- 31. A method for determining the presence of a mammalian G protein-coupled glutamate receptor in a biological sample, which comprises incubating the sample with a monospecific antibody which specifically binds to the receptor under conditions sufficient for immune complex formation and determining therefrom the presence of the immune complexes.
- 32. The method of claim 31, wherein the monospecific antibody is a monoclonal antibody or a purified antiserum.

20

- 33. The method of claim 32, wherein the monospecific antibody is labeled.
- 34. A method for identifying a compound which alters
  25 G protein-coupled glutamate receptor mediated-metabolism, which
  comprises incubating the compound with eukaryotic cells which
  express recombinant mammalian G protein-coupled glutamate
  receptor and determining therefrom the effect of said compound
  on receptor-mediated metabolism in the cells.

30

- 35. The method of claim 34, wherein the compound is incubated with the receptor and ligand.
- 36. The method of claim 35, wherein the ligand is glutamate or quisqualate.

PCT/US91/09422

- 37. The method of claim 34, wherein the eukaryotic cell expresses a human or rodent G protein-coupled glutamate receptor.
- 38. The method of claim 37, wherein inositol phospholipid metabolism in the eukaryotic cell is monitored for alteration by the compound.

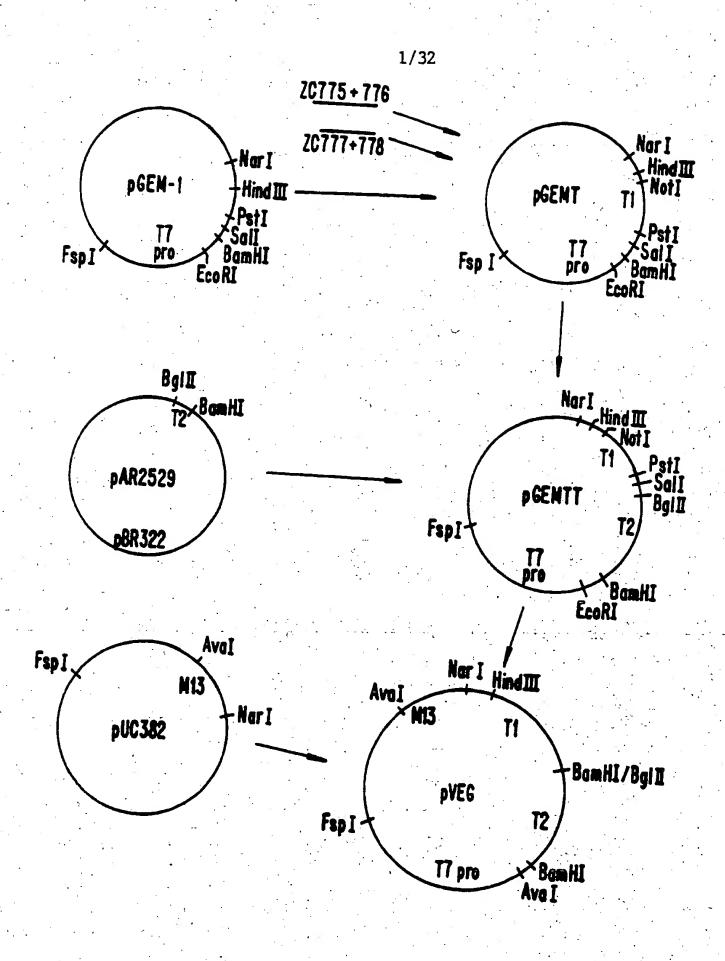


FIG. IA.

2/32

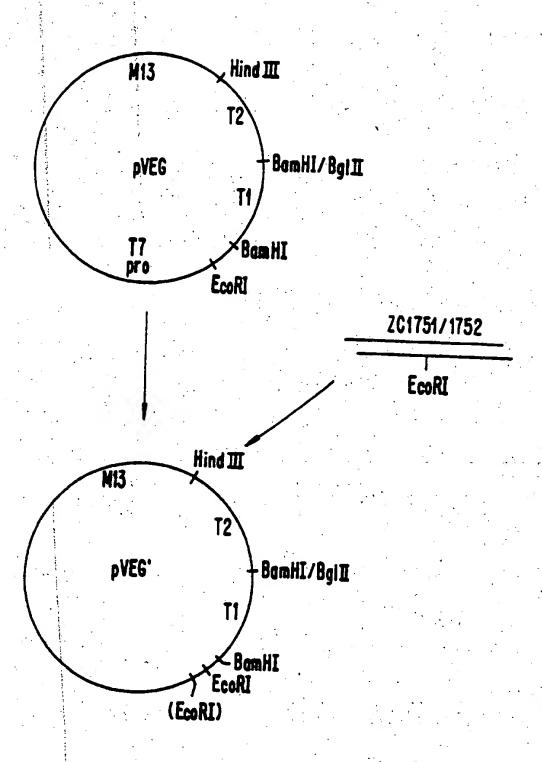


FIG. 1B.

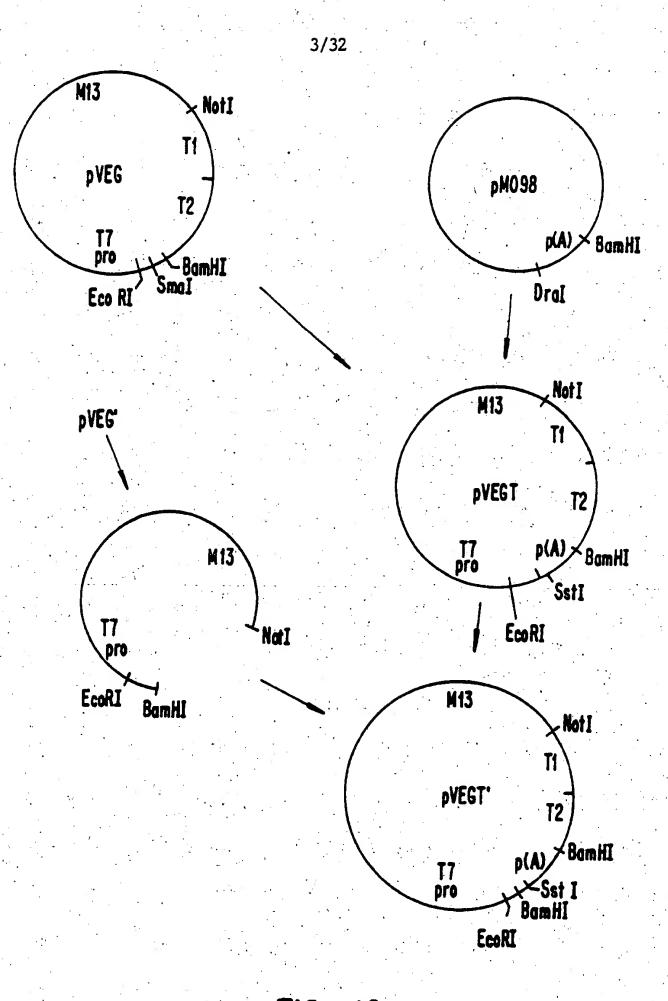
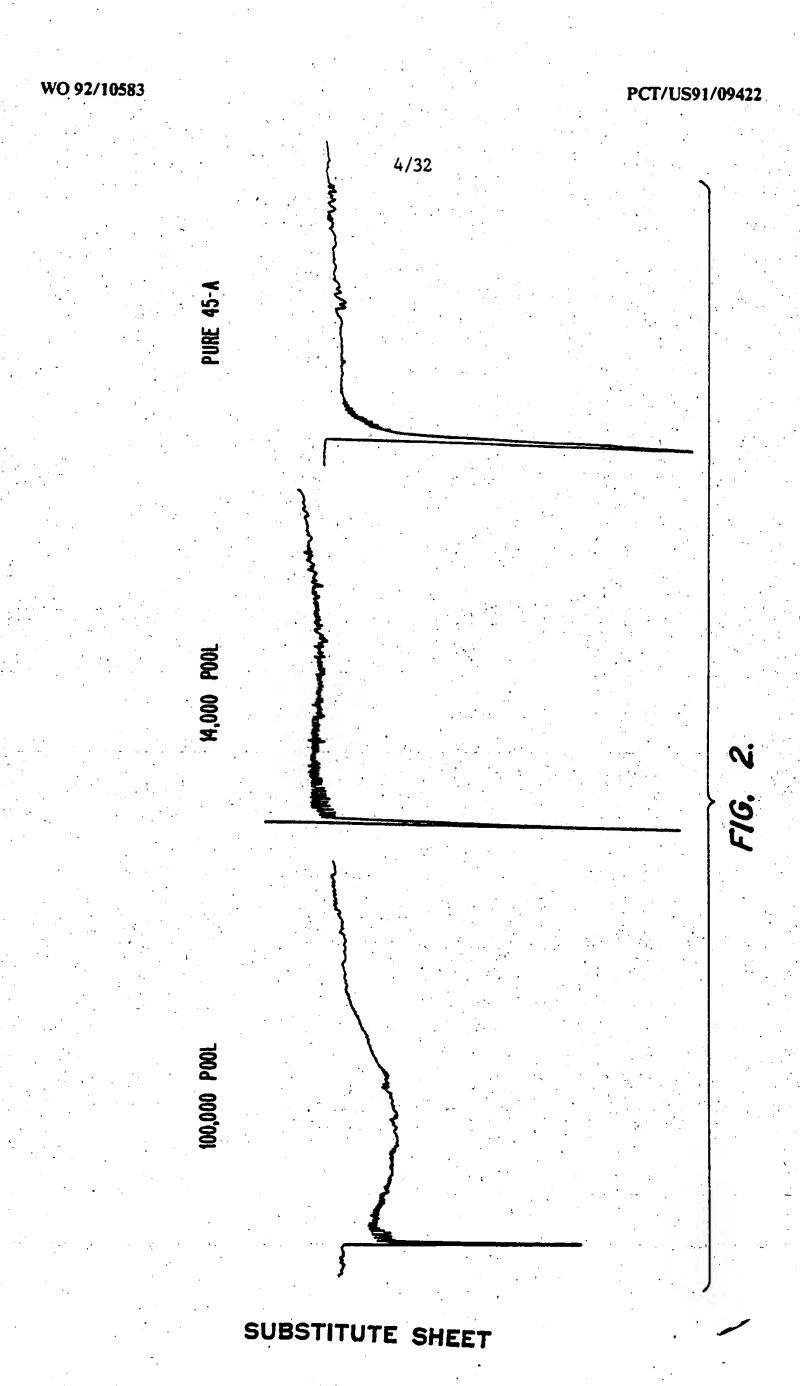
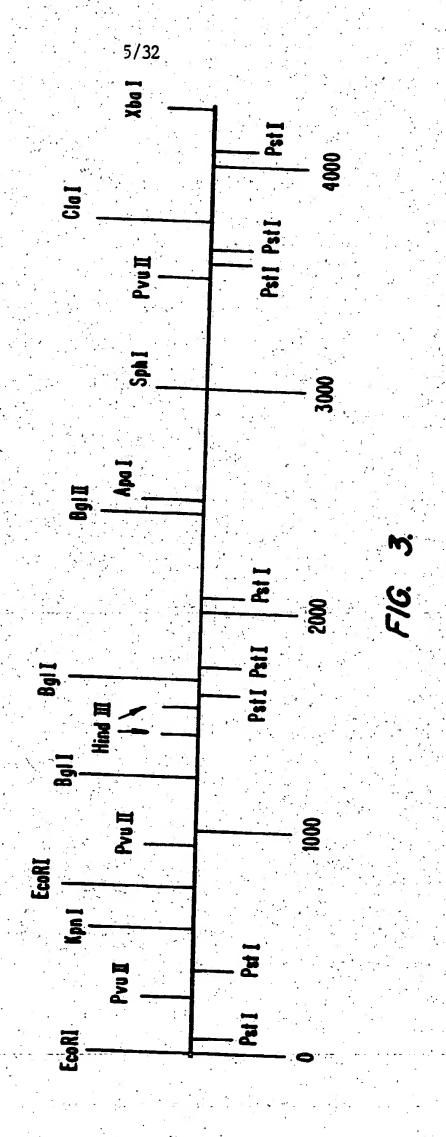
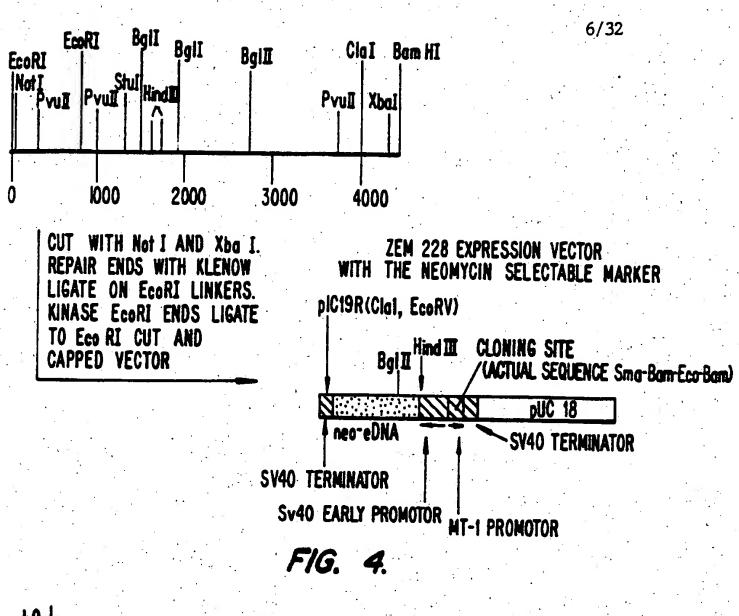


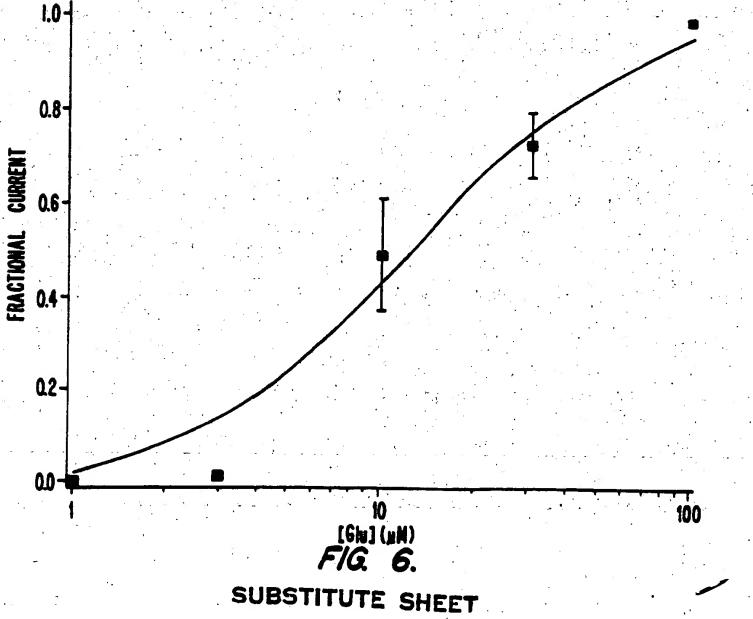
FIG. IC.



WO 92/10583 PCT/US91/09422







7/32 CCGAGAACGG CTGCAGTCCT CTGACCTGAG ACCAATAGCT GTGTCTACCC GGACTCAGCG 120 TCCAGCTCAC CGCCACTAAC GCGCCGCGCA TTGGACACCT GATCCACACA CCTTCGGGCA CCAGTGAAAA ACCGCGACTT GATTTTCTGG AAGAACGCCC CCAGGGTGTG GGAGCGGTCG -240 TGGAGGACCA GCAGGAGGAA GCGGAGGGGA GAGGGCAGT AGTGGAGGCA GAGAAAGCGT TGAACCAGCT GTGTTGGCCG AAGGCACGAA ACGGCAAAAG GCAGCGGTGA GCATCTGTGT GGTTCCCGCT GGGAACCTGC AGGCAGGACC GGCGTGGGAA CGTGGCTGGC CCGCGGTGGA CCGCGTCTTC GCCACA ATG GTC CGG CTC CTC TTG ATT TTC TTC CCA ATG Met Val Arg Leu Leu Ile Phe Phe Pro Met ATC TIT TIG GAG ATG TCC ATT TIG CCC AGG ATG CCT GAC AGA AAA GTA Ile Phe Leu Glu Met Ser Ile Leu Pro Arg Met Pro Asp Arg Lys Val 505 TTG CTG GCA GGT GCC TCG TCC CAG CGC TCC GTG GCG AGA ATG GAC GGA Leu Leu Ala Gly Ala Ser Ser Gln Arg Ser Val Ala Arg Met Asp Gly GAT GTC ATC ATC GGA GCC CTC TTC TCA GTC CAT CAC CAG CCT CCA GCC Asp Val Ile Ile Gly Ala Leu Phe Ser Val His His Gln Pro Pro Ala 601 GAG AAG GTA CCC GAA AGG AAG TGT GGG GAG ATC AGG GAA CAG TAT GGT Glu Lys Val Pro Glu Arg Lys Cys Gly Glu Ile Arg Glu Gln Tyr Gly ATC CAG AGG GTG GAG GCC ATG TTC CAC ACG TTG GAT AAG ATT AAC GCG Ile Gln Arg Val Glu Ala Met Phe His Thr Leu Asp Lys Ile Asn Ala GAC CCG GTG CTC CTG CCC AAC ATC ACT CTG GGC AGT GAG ATC CGG GAC Asp Pro Val Leu Leu Pro Asn Ile Thr Leu Gly Ser Glu Ile Arg Asp 100 TCC TGC TGG CAC TCT TCA GTG GCT CTC GAA CAG AGC ATC GAA TTC ATC Ser Cys Trp His Ser Ser Val Ala Leu Glu Gln Ser Ile Glu Phe 110 120

FIG. 5A.

8/32 AGA GAC TCC CTG ATT TCC ATC CGA GAT GAG AAG GAT GGG CTG AAC CGA Arg Asp Ser Leu Ile Ser Ile Arg Asp Glu Lys Asp Gly Leu Asn Arg 841 TGC CTG CCT GAT GGC CAG ACC CTG CCC CCT GGC AGG ACT AAG AAG CCT Cys Leu Pro Asp Gly Gln Thr Leu Pro Pro Gly Arg Thr Lys Lys Pro 140 145 155 889 ATT GCT GGA GTG ATC GGC CCT GGC TCC AGC TCT GTG GCC ATT CAA GTC Ile Ala Gly Val Ile Gly Pro Gly Ser Ser Val Ala Ile Gln Val 160 CAG AAT CTT CTC CAG CTG TTC GAC ATC CCA CAG ATC GCC TAT TCT GCC Gln Asn Leu Leu Gln Leu Phe Asp Ile Pro Gln Ile Ala Tyr Ser Ala 180 9.85 ACA AGC ATA GAC CTG AGT GAC AAA ACT TTG TAC AAA TAC TTC CTG AGG Thr Ser Ile Asp Leu Ser Asp Lys Thr Leu Tyr Lys Tyr Phe Leu Arg 195 190 200 1033 GTG GTC CCT TCT GAC ACT TTG CAG GCA AGG GCG ATG CTC GAC ATA GTC Val Val Pro Ser Asp Thr Leu Gln Ala Arg Ala Met Leu Asp Ile Val 210 205 1081 AAG CGT TAC AAC TGG ACC TAT GTC TCA GCA GTC CAC ACA GAA GGG AAT Lys Arg Tyr Asn Trp Thr Tyr Val Ser Ala Val His Thr Glu Gly Asn 225 230 1129 TAC GGC GAG AGT GGA ATG GAT GCT TTC AAA GAA CTG GCT GCC CAG GAA Tyr Gly Glu Ser Gly Met Asp Ala Phe Lys Glu Leu Ala Ala Gln Glu 240 245 250 1177 GGC CTC TGC ATC GCA CAC TCG GAC AAA ATC TAC AGC AAT GCT GGC GAG Gly Leu Cys Ile Ala His Ser Asp Lys Ile Tyr Ser Asn Ala Gly Glu 255[.] 1225 AAG AGC TIT GAC CGG CTC CTG CGT AAA CTC CGG GAG CGG CTT CCC AAG Lys Ser Phe Asp Arg Leu Leu Arg Lys Leu Arg Glu Arg Leu Pro Lys 270 280 GCC AGG GTT GTG GTC TGC TTC TGC GAG GGC ATG ACA GTG CGG GGC TTA Ala Arg Val Val Cys Phe Cys Glu Gly Met Thr Val Arg Gly Leu **290**. 285 **295** ·

FIG. 5B.

CTG AGT GCC ATG CGC CGC CTG GGC GTC GTG GGC GAG TTC TCA CTC ATT Leu Ser Ala Met Arg Arg Leu Gly Val Val Gly Glu Phe Ser Leu Ile 300 305 13.69 GGA AGT GAT GGA TGG GCA GAC AGA GAT GAA GTC ATC GAA GGC TAT GAG Gly Ser Asp Gly Trp Ala Asp Arg Asp Glu Val Ile Glu Gly Tyr Glu 325 GTG GAA GCC AAC GGA GGG ATC ACA ATA AAG CTT CAG TCT CCA GAG GTC Val Glu Ala Asn Gly Gly Ile Thr Ile Lys Leu Gln Ser Pro Glu Val 340 AGG TCA TTT GAT GAC TAC TTC CTG AAG CTG AGG CTG GAC ACC AAC ACA Arg Ser Phe Asp Asp Tyr Phe Leu Lys Leu Arg Leu Asp Thr Asn Thr AGG AAT CCT TGG TTC CCT GAG TTC TGG CAA CAT CGC TTC CAG TGT CGC Arg Asn Pro Trp Phe Pro Glu Phe Trp Gln His Arg Phe Gln Cys Arg 365 CTA CCT GGA CAC CTC TTG GAA AAC CCC AAC TTT AAG AAA GTG TGC ACA Leu Pro Gly His Leu Leu Glu Asn Pro Asn Phe Lys Lys Val Cys Thr 380 395 1609 GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AAA ATG GGA Gly Asn Glu Ser Leu Glu Glu Asn Tyr Val Gln Asp Ser Lys Met Gly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CAG AAC ATG Phe Val Ile Asn Ala Ile Tyr Ala Met Ala His Gly Leu Gin Asn Met 415 420 425 1705 CAC CAT GCT CTG TGT CCC GGC CAT GTG GGC CTG TGT GAT GCT ATG AAA His His Ala Leu Cys Pro Gly His Val Gly Leu Cys Asp Ala Met Lys 430 435 CCC ATT GAT GGC AGG AAG CTC CTG GAT TTC CTC ATC AAA TCC TCT TTT Pro Ile Asp Gly Arg Lys Leu Leu Asp Phe Leu Ile Lys Ser Ser Phe 445 450 · 1801 GTC GGA GTG TCT GGA GAG GAG GTG TGG TTC GAT GAG AAG GGG GAT GCT Val Gly Val Ser Gly Glu Glu Val Trp Phe Asp Glu Lys Gly Asp Ala 460° 465 470 - 475

FIG. 5C.

							٠,				•					
	CCC Pro	GGA Gly	AGG Arg	TAT Tyr	GAC Asp 480	ATT	ATG Met	10/32 AAT Asn	CTG	CAG Gln 485	TAC Tyr	ACA Thr	GAA Glu	GCT Ala	· AAT	1849 CGC Arg
•	TAT Tyr	GAC Asp	TAT Tyr	GTC Val 495	CAC His	GTG Val	GGG Gly	ACC Thr	TGG Trp 500	CAT His	GAA Glu	GGA Gly	GTG Val	CTG Leu 505	AAT Asn	1897 ATT Ile
	GAT Asp	GAT Asp	TAC Tyr 510	AAA Lys	ATC	CAG Gln	ATG Met	AAC Asn 515	AAA Lys	AGC Ser	GGA Gly	ATG Met	GTA Val 520	CGA Arg	TCT	1945 GTG Val
	TGC Cys	AGT Ser 525	GAG G1u	CCT Pro	TGC Cys	TTA Leu	AAG Lys 530	GGT Gly	CAG Gln	ATT	AAG Lys	GTC Val 535	ATA Ile	CGG Arg	AAA	1993 GGA Gly
	GAA G1u 540	GTG Val	AGC Ser	TGC Cys	TGC Cys	TGG Trp 545	ATC Ile	TGC Cys	ACG Thr	GCC Ala	TGC Cys 550	AAA Lys	GAG G1 u	AAT Asn	GAG	2041 TTT Phe 555
	GTG Val	CAG G1n	GAC Asp	Glu	TTC Phe 560	ACC Thr	TGC Cys	AGA Arg	GCC Ala	TGT Cys 565	GAC Asp	CTG Leu	GGG Gly	TGG Trp	TGG	2089 CCC Pro
	AAC Asn	GCA Ala	GAG G1u	CTC Leu 575	Thr	GGC Gly	TGT Cys	GAG G1u	CCC Pro 580	ATT Ile	CCT Pro	GTC Val	CGT Arg		CTT ² Leu	
	TGG Trp		GAC Asp 590	ATA Ile	GAA Glu	TCT Ser	ATC Ile	ATA Ile 595	GCC Ala	ATC Ile	GCC Ala	TTT Phe		TGC Cys	CTG	GGC Gly
	ATC		GTG Val	ACG Thr	CTG Leu	TTT Phe	GTC Val 610	ACC Thr	CTC Leu	ATC Ile	TTC Phe	GTT Val 615	CTG Leu	TAC Tyr	CGG	233 GAC Asp
•					AAA Lys							TGC Cys			ATT Ile	
												ACC Thr		ATC Ile	GCC	2329 AAA Lys

FIG. 5D.

CCCGGGCTCC CGGCAGTGCG AGCAGCTAAG GGCTGGCCGC CGCCTCCCTG AGCTCCCCG 120 GMGCAGCCGA CCCCTGGTCG CGGCGTTCAC CTCGCCGATG CGCGGTTGGT AGGAGTGACC GGAGCCATTC TCTCCTCGTT GATAAGATTC CCTACCAGGA TAGGAGCCTA TCTCCCTTTY CACAGCAGGA CACAGAAATC TGGCCTTCAG TACTTTGGGA AAAGGATCTG AGACCTCCTG GAGCTCTGAC CACTGGCTGT CATCTGTGGC TCTGGCCTGT GTGGGCCACT GAGCTCTACT CAAACATTAA AGAGGAGGAG GGGAGATCTG TGGAATGGGC CACCCCGTTG GCCTGCTGCA TTACTGAACC TGCGCTGTCC ACACGTGCCC AGATCATGGG ACCCAGGGCC TGCTAGGGCT AGGAGCGGGG CCCAGTATTC ATGGGTCTCT AGGCCTTTCC GAA ATG TCC GGG AAG Met Ser Gly Lys GGA GGC TGG GCC TGG TGG GCC CGG CTG CCC CTC TGC CTA CTC CTC Gly Gly Trp Ala Trp Trp Ala Arg Leu Pro Leu Cys Leu Leu Leu AGC CTT TAT GCC CCC TGG GTG CCT TCA TCC TTG GGA AAG CCC AAG GGT Ser Leu Tyr Ala Pro Trp Val Pro Ser Ser Leu Gly Lys Pro Lys Gly CAC CCC CAC ATG AAC TCT ATC CGA ATT GAC GGG GAC ATC ACA CTG GGA His Pro His Met Asn Ser Ile Arg Ile Asp Gly Asp Ile Thr Leu Gly GGC CTG TTT CCC GTC CAC GGC CGT GGC TCT GAG GGT AAG GCC TGC GGG Gly Leu Phe Pro Val His Gly Arg Gly Ser Glu Gly Lys Ala Cys Gly GAG CTG AAG AAG GAG AAA GGC ATC CAC CGC CTG GAG GCC ATG CTG TTT Glu Leu Lys Lys Glu Lys Gly Ile His Arg Leu Glu Ala Met Leu Phe **70**: 763 GCC CTG GAC CGC ATC AAC AAT GAC CCG GAC CTA CTG CCC AAC ATC ACG Ala Leu Asp Arg Ile Asn Asn Asp Pro Asp Leu Leu Pro Asn Ile Thr 90 85 100

FIG. 8A.

24/32

TTG GGC GCC CGC ATT CTG GAC ACC TGC TCG AGG GAC ACC CAC GCC CTG Leu Gly Ala Arg Ile Leu Asp Thr Cys Ser Arg Asp Thr His Ala Leu GAG CAG TCA CTG ACC TTT GTG CGG GCG CTC ATC GAG AAG GAC GGC ACG Glu Gln Ser Leu Thr Phe Val Arg Ala Leu Ile Glu Lys Asp Gly Thr GAG GTC CGC TGC GGC AGG CGG GGC CCC ATC ATC ACC AAG CCC GAA Glu Val Arg Cys Gly Arg Arg Gly Pro Pro Ile Ile Thr Lys Pro Glu CGA GTG GTG GGT GTC ATT GGA GCT TCG GGG AGC TCC GTC TCG ATC ATG Arg Val Val Gly Val Ile Gly Ala Ser Gly Ser Ser Val Ser Ile Met 150 GTG GCC AAC ATC CTC CGC CTC TTC AAG ATC CCT CAG ATC AGC TAT GCC Val Ala Asn Ile Leu Arg Leu Phe Lys Ile Pro Gln Ile Ser Tyr Ala TCC ACG GCC CCT GAC TTG AGT GAC AAC AGC CGC TAT GAC TTC TTC TCC Ser Thr Ala Pro Asp Leu Ser Asp Asn Ser Arg Tyr Asp Phe Phe Ser 185 CGG GTG GTG CCC TCA GAC ACA TAC CAG GCC CAG GCC ATG GTG GAT ATT Arg Val Val Pro Ser Asp Thr Tyr Gln Ala Gln Ala Met Val Asp Ile 200 205 GTC CGA GCC CTC AAG TGG AAC TAT GTG TCC ACA CTG GCC TCA GAG GGC Val Arg Ala Leu Lys Trp Asn Tyr Val Ser Thr Leu Ala Ser Glu Gly 215 AGC TAC GGT GAG AGT GGT GTG GAG GCC TTT ATC CAG AAG TCC CGA GAG Ser Tyr Gly Glu Ser Gly Val Glu Ala Phe Ile Gln Lys Ser Arg Glu 230 235 AAC GGA GGT GTG TGC ATT GCC CAG TCG GTG AAG ATT CCA CGG GAA CCC Asn Gly Gly Val Cys Ile Ala Gln Ser Val Lys Ile Pro Arg Glu Pro 260 245 AAG ACG GGG GAG TTC GAC AAG ATC ATC AAA CGC CTA CTG GAA ACA TCC Lys Thr Gly Glu Phe Asp Lys Ile Ile Lys Arg Leu Leu Glu Thr Ser 265

FIG. 8B.

				œ				0= /0/								
			AGG Arg	GGT Gly 280	ATC Ile	ATC Ile	ATC Ile	25/32 TTT Phe	GCC	Asn.	GAG G1 u	GAT Asp	GAC Asp	ATC Ile 290	AGG Arg	
												GGC Gly			TTT Phe	
	ATG Met														CGC Arg	
•	GAG Glu 325											CCC Pro			ATG	483 TCT Ser 340
	GTT Val	CGA Arg	GGG Gly	TTC Phe	GAC Asp 345	CGA Arg	TAC Tyr	TTC Phe	TCC Ser	AGC Ser 350	CGC Arg	ACG Thr	CTG Leu	GAC Asp	AAC Asn 355	.531 AAC Asn
	AGG Arg	CGC Arg	AAC Asn	ATC Ile 360	TGG Trp	TTT Phe	GCC Ala	GAG G1u	TTC Phe 365	TGG Trp	GAG G1u	GAC Asp	AAC Asn	TTC Phe 370	CAT His	579 TGC Cys
	1	TTG Leu										CAC His				627 TGC Cys
	Thr											TAT Tyr 400			GAG	675 GGG G1y
	AAG Lys 405											ATG Met			GCG	723 CTG Leu 420
	CAC His	GCC Ala	ATG Met	CAC His	CGT Arg 425	GAC Asp	CTG Leu	TGT Cys	CCC	GGC Gly 430	CGC Arg	GTA Val	GGA Gly	CTC Leu	TGC	771 CCT Pro
						GAT Asp						AAG Lys			AGG Arg	_
	,	••		_			٠.									

FIG. 8C.

26/32 GTC AAC TTC TCA GGC ATT GCG GGG AAC CCT GTA ACC TTC AAT GAG AAC Val Asn Phe Ser Gly Ile Ala Gly Asn Pro Val Thr Phe Asn Glu Asn 460 GGA GAC GCA CCG GGG CGC TAC GAC ATC TAC CAG TAC CAA CTG CGC AAT Gly Asp Ala Pro Gly Arg Tyr Asp Ile Tyr Gln Tyr Gln Leu Arg Asn GGC TCG GCC GAG TAC AAG GTC ATC GGC TCG TGG ACA GAC CAC CTG CAC Gly Ser Ala Glu Tyr Lys Val Ile Gly Ser Trp Thr Asp His Leu His 490 CTC AGA ATA GAG CGG ATG CAG TGG CCA GGG AGT GGC CAG CAG CTG CCG Leu Arg Ile Glu Arg Met Gln Trp Pro Gly Ser Gly Gln Gln Leu Pro 505 510 CGC TCC ATC TGC AGT CTG CCC TGC CAG CCC GGG GAG CGA AAG AAG ACT Arg Ser Ile Cys Ser Leu Pro Cys Gln Pro Gly Glu Arg Lys Lys Thr GTG AAG GGC ATG GCT TGC TGC TGC CAC TGC GAG CCC TGC ACC GGG TAC Val Lys Gly Met Ala Cys Cys Trp His Cys Glu Pro Cys Thr Gly Tyr 535 CAG TAC CAA GTG GAC CGC TAC ACC TGT AAG ACC TGC CCC TAC GAC ATG Gln Tyr Gln Val Asp Arg Tyr Thr Cys Lys Thr Cys Pro Tyr Asp Met 550 CGG CCC ACA GAG AAC CGC ACG AGC TGC CAG CCC ATC CCC ATC GTC AAG Arg Pro Thr Glu Asn Arg Thr Ser Cys Gln Pro Ile Pro Ile Val Lys 570 575 × 565 TTG GAG TGG GAC TCG CCG TGG GCC GTG CTG CCC CTC TTC CTG GCC GTG Leu Glu Trp Asp Ser Pro Trp Ala Val Leu Pro Leu Phe Leu Ala Val **585** GTG GGC ATC GCC GCC ACG CTG TTC GTG GTG GTC ACG TTT GTG CGC TAC Val Gly Ile Ala Ala Thr Leu Phe Val Val Val Thr Phe Val Arg Tyr 600 605. 610 AAC GAT ACC CCC ATC GTC AAG GCC TCG GGC CGG GAG CTG AGC TAC GTG Asn Asp Thr Pro Ile Val Lys Ala Ser Gly Arg Glu Leu Ser Tyr Val 615 620

FIG. 8D.

27/32 CTG CTG GCG GGC ATC ITT CTG IGC IAC GCC ACT ACC ITC CTC ATG ATC Leu Leu Ala Gly Ile Phe Leu Cys Tyr Ala Thr Thr Phe Leu Met Ile 630 GCA GAG CCG GAC CTG GGG ACC TGT TCG CTC CGC CGC ATC TTC CTA GGG Ala Glu Pro Asp Leu Gly Thr Cys Ser Leu Arg Arg Ile Phe Leu Gly CTC GGC ATG AGC ATC AGC IAC GCG GCC CTG CTG ACC AAG ACC AAC CGC Leu Gly Met Ser Ile Ser Tyr Ala Ala Leu Leu Thr Lys Thr Asn Arg 665 ATT TAC CGC ATC TTT GAG CAG GGC AAA CGG TCG GTC AGT GCC CCG CGT Ile Tyr Arg Ile Phe Glu Gln Gly Lys Arg Ser Val Ser Ala Pro Arg 680 685 TTC ATC AGC CCG GCC TCG CAG CTG GCC ATC ACC TTC ATC CTC ATC TCC Phe Ile Ser Pro Ala Ser Gln Leu Ala Ile Thr Phe Ile Leu Ile Ser 700 · 705 CTG CAG CTG CTC GGC ATC TGC GTG TGG TTC GTG GTG GAC CCC TCC CAC Leu Gln Leu Leu Gly Ile Cys Val Trp Phe Val Val Asp Pro Ser His 710 TCG GTG GTG GAC TTC CAG GAC CAA CGG ACA CTT GAC CCC CGC TTT GCC Ser Val Val Asp Phe Gln Asp Gln Arg Thr Leu Asp Pro Arg Phe Ala 730 AGG GGC GTG CTC AAG TGC GAC ATC TCG GAC CTG TCC CTC ATC TGC CTG Arg Gly Val Leu Lys Cys Asp Ile Ser Asp Leu Ser Leu Ile Cys Leu 750 CTG GGC TAC AGC ATG CTG CTG ATG GTC ACG TGT ACT GTG TAC GCC ATC Leu Gly Tyr Ser Met Leu Leu Met Val Thr Cys Thr Val Tyr Ala Ile 765 AAG ACC CGA GGC GTG CCC GAG ACC TTC AAC GAG GCC AAG CCC ATC GGC Lys Thr Arg Gly Val Pro Glu Thr Phe Asn Glu Ala Lys Pro Ile Gly 7**80**. ... TTC ACC ATG TAC ACC ACC TGC ATT GTC TGG CTG GCC TTC ATC CCC ATC Phe Thr Met Tyr Thr Thr Cys Ile Val Trp Leu Ala Phe Ile Pro Ile 790 800 **790** 

FIG. 8E.

2923 TTT TTT GGC ACC TCA CAG TCA GCC GAC AAG CTG TAC ATC CAG ACA ACC Phe Phe Gly Thr Ser Gln Ser Ala Asp Lys Leu Tyr Ile Gln Thr Thr ~ 810 820 2971 ACA CTG ACG GTC TCC GTG AGT CTG AGC GCT TCA GTG TCC CTG GGG ATG Thr Leu Thr Val Ser Val Ser Leu Ser Ala Ser Val Ser Leu Gly Met 825 830 3019 CTC TAC ATG CCC AAA GTC TAC ATC ATC CTC TTC CAC CCG GAG CAG AAC Leu Tyr Met Pro Lys Val Tyr Ile Ile Leu Phe His Pro Glu Gln Asn 840 845 3067 GTG CCC AAG CGC AAG CGC AGT CTC AAA GCC GTG GTC ACC GCC GCC ACC Val Pro Lys Arg Lys Arg Ser Leu Lys Ala Val Val Thr Ala Ala Thr 860 ATG TCC AAC AAG ITC ACA CAG AAG GGC AAC TTC AGG CCC AAT GGG GAA Met Ser Asn Lys Phe Thr Gln Lys Gly Asn Phe Arg Pro Asn Gly Glu 875 GCC AAA TCA GAG CTG TGT GAG AAC CTG GAG ACC CCA GCG CTG GCT ACC Ala Lys Ser Glu Leu Cys Glu Asn Leu Glu Thr Pro Ala Leu Ala Thr 890 885 895 900 AAA CAG ACC TAC GTC ACC TAC ACC CAT GCC ATC TAGCCGGGCC Lys Gln Thr Tyr Val Thr Tyr Thr Asn His Ala Ile 905 · 3269

GCGGAGCCAA GCAGGCTAAG GAGCCACAAC CTCTGAGGAT GGCACATTGG GCCAGGGCCG

TTCCCGAGGG CCCTGCCGAT GTCTGCCCGC CTCCCGGGCA TCCACGAATG TGGCTTGGTG

CTGAGGACAG TAGAGACCCC GGCCATCACT GCTGGGCAAG CCGTGGTGGG CAACCAGAGG

AGGCCGAGTG GCTGGGGCAG TTCCAGGTTA TGCCACACAC AGGTCTTCCT TCTGGACCAC

TGTTGGCCCA GCCCCAAAGC ACAGGGGCTC GGTCTCCAGA GCCCAGCCCT GGCTTCCTCT

CCTTCCTCCT GCCTCCGTCT GTCCTGTGGG TGACCCCGGT TGGTCCCTGC CCCGTCTTTA

CGTTTCTCTT CCGTCTTTGC TCTGCATGTG TTGTCTGTTT GGGCCCTCTG CTTCCATATT

FIG. 8F.

TTTCCATTCT GCTCCTGGCC TTCCCCTGCC ATCTGCCCTG CCCCCTGCCC CTCCTCCTG

AGCTGCCCCA TCCCCGCCAT CATTTTCTCT TCTGTTCCCC CTCGATCTCA TTTCCTACCA

GCCTTCCCCC TACTTGGCTT CATCCACCAA CTCTTTCACC ACGTTGCAAA AGAGAAAAAA

AAAGGGGGGG GGGAATCACC CCCTACAAAA AAGCCCAAAC AAAAACTAAT CTTGAGTGTG

TTTCGAAGTG CTGCGTCCTC CTGGTGGCCT GTGTGTCCCT GTGCCTGCAG CCTGTCTGCC

CGCCCTACCC GTCTGCCGTG TGTCCTGCCC CCCCCGCCTG CCCGCCTTGC CCTTCCTGCT

AACGACACGG AGTTCAGTGC CTGGGTGTTT GGTGATGGTC TCTGATGTGT AGCATGTCTG

TTTTTATACC GAGAACATTT CTAATAAAGA TAAACACATG GTTTTGC

FIG. 8G.

CCCAACATCA CGTTGGGCGC CCGCATTCTG GACACCTGCT CGAGGGACAC CCACGCCCTG GAGCAGTCAC TGACCTTTGT GCGGGCGCTC ATCGAGAAGG ACGGCACGGA GGTCCGCTGC 180 GGCAGGCGGG GCCCGCCCAT CATCACCAAG CCCGAACGAG TGGTGGGTGT CATTGGAGCT TCGGGGAGCT CCGTCTCGAT CATGGTGGCC AACATCCTCC GCCTCTTCAA GATCCCTCAG ATCAGCTATG CCTCCACGGC CCCTGACTTG AGTGACAACA GCCGCTATGA CTTCTTCTCC CGGGTGGTGC CCTCAGACAC ATACCAGGCC CAGGCCATGG TGGATATTGT CCGAGCCCTC AAGTGGAACT ATGTGTCCAC ACTGGCCTCA GAGGGCAGCT ACGGTGAGAG TGGTGTGGAG GCCTTTATCC AGAAGTCCCG AGAGAACGGA GGTGTGTGCA TTGCCCAGTC GGTGAAGATT CCACGGGAAC CCAAGACGGG GGAGTTCGAC AAGATCATCA AACGCCTACT GGAAACATCC AATGCCAGGG GTATCATCAT CTTTGCCAAC GAGGATGACA TCAGGAGGGT GTTGGAGGCA GCTCGCAGGG CCAACCAGAC CGGCCACTTC TTTTGGATGG GTTCTGATAG CTGGGGCTCC AAGAGTGCCC CTGTGCTGCG CCTTGAGGAG GTGGCCGAGG GCGCAGTCAC CATTCTCCCC AAGAGGATGT CTGTTCGAGG GTTCGACCGA TACTTCTCCA GCCGCACGCT GGACAACAAC AGGCGCAACA TCTGGTTTGC CGAGTTCTGG GAGGACAACT TCCATTGCAA GTTGAGCCGC CACGCGCTCA AGAAGGGAAG CCACATCAAG AAGTGCACCA ACCGAGAGCG CATCGGGCAG GACTCGGCCT ATGAGCAGGA GGGGAAGGTG CAGTTCGTGA TTGACGCTGT GTACGCCATG 1020 GGCCACGCGC TGCACGCCAT GCACCGTGAC CTGTGTCCCG GCCGCGTAGG ACTCTGCCCT 1080 CGCATGGACC CCGTGGATGG CACCCAGCTG CTTAAGTACA TCAGGAACGT CAACTTCTCA 1140 GGCATTGCGG GGAACCCTGT AACCTTCAAT GAGAACGGAG ACGCACCGGG GCGCTACGAC

FIG. 9A.

31/32 1200 ATCTACCAGT ACCAACTGCG CAATGGCTCG GCCGAGTACA AGGTCATCGG CTCGTGGACA 1260 GACCACCTGC ACCTCAGAAT AGAGCGGATG CAGTGGCCAG GGAGTGGCCA GCAGCTGCCG CGCTCCATCT GCAGTCTGCC CTGCCAGCCC GGGGAGCGAA AGAAGACTGT GAAGGGCATG 1380 GCTTGCTGCT GGCACTGCGA GCCCTGCACC GGGTACCAGT ACCAAGTGGA CCGCTACĂCC TGTAAGACCT GCCCCTACGA CATGCGGCCC ACAGAGAACC GCACGAGCTG CCAGCCCATC 1500 CCCATCGTCA AGTTGGAGTG GGACTCGCCG TGGGCCGTGC TGCCCCTCTT CCTGGCCGTG GTGGGCATCG CCGCCACGCT GTTCGTGGTG GTCACGTTTG TGCGCTACAA CGATACCCCC 1620 ATCGTCAAGG CCTCGGGCCG GGAGCTGAGC TACGTGCTGC TGGCGGGCAT CTTTCTGTGC 1680 TACGCCACTA CCTTCCTCAT GATCGCAGAG CCGGACCTGG GGACCTGTTC GCTCCGCCGC ATCTTCCTAG GGCTCGGCAT GAGCATCAGC TACGCGGCCC TGCTGACCAA GACCAACCGC ATTTACCGCA TCTTTGAGCA GGGCAAACGG TCGGTCAGTG CCCCGCGTTT CATCAGCCCG GCCTCGCAGC TGGCCATCAC CTTCATCCTC ATCTCCCTGC AGCTGCTCGG CATCTGCGTG 1920 TGGTTCGTGG TGGACCCCTC CCACTCGGTG GTGGACTTCC AGGACCAACG GACACTTGAC CCCCCCTTTG CCAGGGGCGT GCTCAAGTGC GACATCTCGG ACCTGTCCCT CATCTGCCTG 2040 CTGGGCTACA GCATGCTGCT GATGGTCACG TGTACTGTGT ACGCCATCAA GACCCGAGGC GTGCCCGAGA CCTTCAACGA GGCCAAGCCC ATCGGCTTCA CCATGTACAC CACCTGCATT 2160 GTCTGGCTGG CCTTCATCCC CATCTTTTT GGCACCTCAC AGTCAGCCGA CAAGCTGTĂC 2220 ATCCAGACAA CCACACTGAC GGTCTCCGTG AGTCTGAGCG CTTCAGTGTC CCTGGGGATG 2280 CTCTACATGC CCAAAGTCTA CATCATCCTC TTCCATATTT TTCCATTCTG CTCCTGGCCT

FIG. 9B.

32/32

TCCCCTGCCA TCTGCCCTGC CCCCTGCCCC TCCTCCCTGA GCTGCCCCAT CCCCGCCATC

ATTTTCTCTT CTGTTCCCCC TCGATCTCAT TTCCTACCAG CCTTCCCCCT ACTTGGCTTC

CTCCACCAAC TCTTTCACCA CGTTGC

FIG 9C

### INTERNATIONAL SEARCH REPORT

International Application No. PCT/US91/09422

	CLASSIFICATION F SUBJECT MATTER (if several classification symbols apply, indicate all) ³ .ccording to International Patent Classification (IPC) or to both National Classification and IPC												
1			• .	th National Classificat	ion and IPC								
US CL	•	e See Attached She 435/69.1, 240.	2, 320.1; 53	0/350, 351, 387;	536/27.	· · ·							
II. FIELI	DS SEARC		dinimum Docum	mentation Searched	l :								
Classificati	on System	<del></del>	<del></del>	Classification Symbo									
Classificati	· ·		·	Grassification Gymbi	7.0								
U.S.		US CL : 435/6	9.1, 240.2	, 320.1; 530/3	50, 351, 3	387; 536/27.							
					7 .	- · · ·							
				l other than Minimum ments are included in									
cas, c	online,	aps		н	• • •	- Re-							
·	:			ti de		*							
III. DOC	uments (	CONSIDERED TO BE RE	LEVANT 14		•								
Category*	Citatio	n of Document,16 with ind	cation, where ap	propriate, of the relevant	passages ¹⁷	Relevant to Claim No. 18							
х/ү	et al., "A new type of glutamate receptor linked to inositol phospholipid metabolism", pages 531-533, see the entire document.												
x/y													
Y	et al. of the	, Volume 342, i , "Cloning by f glutamate rece tire document.	unctional	expression of	a member	1-3 and 6-30							
x,p	al.,	, Volume 349, i "sequence and ate receptor", p	expressio	n of a meta	botropic	1-3, 6-30							
- Sc			*		-36	50 P							
1	0				,								
•			E			* * * * * * * * * * * * * * * * * * * *							
	er		_	*									
			-	•									
		*	× 1,*										
"A" doct not inter inter or vanot "C" doct or o or o or o doct or o doct but i	ument defin considered er document mational fili ument whice which is cit ther citation ument refer ther means ument publi	h may throw doubts on p ed to establish the public or other special reason (a ring to an oral disclosure, a shed prior to the Internation of priority date claimed	or after the nortry claim(s) cation date of specified) use, exhibition	date or priority application but theory underlyi "X" document of invention cann considered to i "Y" document of invention can inventive step one or more oth being obvious	y date and no cited to unde ng the invention particular rel not be consider not be consider the documer such documer such documer such documer such sein skill	evance; the claimed ed novel or cannot be nitive step evance; the claimed dered to involve an ment is combined with ents, such combination							
		ompletion of the interna	tional Search ²	Date of Mailing of th	s International	Search Report 2							
	MARCH	1			33AD 100								
		ng Authority ¹	· / (1	Signature of Authoria		$\mathcal{G}(\mathcal{F})$							
	A/US	· ·		Gian Wang P	JIV	April fr							

# FURTHER INFORMATION CONTINUED FROM PREVIOUS SHEETS (Not for publication)

I. CLASSIFICATION OF SUBJECT MATTER: IPC (5):

C12P 21/06; C12N 5/00, 15/00; C07H 15/12, 17/00; C07K 3/00, 13/00, 15/00, 17/00; A61K 35/14.

#### VI. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING This ISA found multiple inventions as follows:

#### Detailed reasons for holding lack of invention

The claims of the three groups have the characteristics of three distinct inventive concepts. Groups I-III are separate and distinct inventions, and require materially different considerations and searches.

#### Itemized summary of claims groupings

- I. Claims 1-3 and 6-30 are drawn to a method for producing a mammalian G protein by using its encoding sequence, classified in Class 435, subclass 69.1, 240.2; Class 530, subclass 387; Class 536. Subclass 27.
- 530, subclass 387; Class 536, Subclass 27.

  II. Claims 4-5 and 31-33 are drawn to a method for determining the presence of a mammalian G protein by using monoclonal antibody, classified in Class 435, subclass 7,21; Class 424, subclass 85.8.
- III. Claims 34-38 are drawn to a method for identifying a compound, classified in Class 435, subclass 4.

the Amersham random-priming kit (Amersham, Arlington Hts, IL). Duplicate lifts were prepar d from the plates, and the filters were hybridized with the probes in 50% formamide at 37°C. After an overnight hybridization, the filters were washed in 2x SSC + 0.1% SDS at 50°C. Positive plaques were isolated by several rounds of dilution plating and repeated screening with the random-primed probes.

10

15

5

#### Table 3

#### NZY Agar

To 950 ml of deionized water, add:

10 g NZ amine: Casein hydrolysate enzymatic (ICN Biochemicals)

5 g NaCl

5 g bacto-yeast extract

1 g casamino acids

2 g  $MgSO_4$  '  $7H_2O$ 

20

Shake until the solutes have dissolved, Adjust to pH 7.0 with 5 N NaOH (approximately 0.2 ml). Adjust the volume of the solution to 1 liter with deionized  $\rm H_2O$ . Sterilize by autoclaving for 20 minutes.

25

#### 20x SSC

Dissolve 175.3 g NaCl and 88.2 g sodium citrate in 800 ml  $H_2O$ . Adjust the pH to 7.0 with a few drops of 10 N NaOH. Adjust the volume to 1 liter with  $H_2O$ . Sterilize by autoclaving.

30

35

Plasmid DNA was prepared from positive plaques using the Bluescript system (Stratagene Cloning Systems). The plasmid DNA was subjected to restriction analysis and Southern blot analysis (Sambrook et al., ibid., which is incorporated her in by reference). Two clon s, SN23, derived from the total rat brain library, and SR2, derived from the rat cerebellum library, were id ntified

10

15

20

25

30

35

as being different than the 45-A cl ne and were sequenc d. Sequenc analysis showed that they represented two additional subtypes. SN23 encodes subtype 1b, which contains an additional 85 bp exon that encodes a new stretch of 20 amino acids and a stop codon in the intracellular domain, is 292 amino acids shorter than the 45-A clone. The nucleotide sequence and deduced amino acid sequence of clone SN23 are shown in Fig. 7. SR2 was found to contain a partial cDNA sequence encoding subtype 2a, which is a novel sequence that shares a 42% homology to the transmembrane domains and extracellular domain of the 45-A clone.

A complete subtype 2a clone was obtained by rescreening both libraries as described above with the radiolabeled 1.3 kb Pst I fragment from clone 45-A and a radiolabeled 1.4 kb Eco RI-Pvu II fragment from SR2. Two additional clones were obtained. SN30, derived from the total rat brain library, contained the entire subtype 2a coding sequence. The nucleotide sequence and deduced amino acid sequence of clone SN30 are shown in Fig. 8. SR13, derived from the rat cerebellum library, contained an incomplete sequence of a new receptor subtype, 2b. Sequence analysis of SR13 showed that the coding sequence was incomplete at the 3' end and was virtually identical to the SN30 sequence except that it contained a 610 base pair deletion within the 3' terminus of SN30. The DNA sequence of the cDNA insert in clone SR13 is shown in Figure 9.

The complete 3' end of the subtype 2a clone was generated using PCR amplification and an oligonucleotide containing a sequence unique to SR13 (ZC4520, Table 4) and an oligonucleotide corresponding to a sequence near the 3' end of the 3' non-translated region of SN30 (ZC4519, Table 4). DNA was prepared from plate lysates of the original plating of each library. Each plate produced a pool of clones. For the PCR reactions, ten nanograms from each library and 100 pmol of ach oligonucleotide w re combined in a reaction volume of 50

μl containing 50 mM KCl, 10 mM Tris-HCl pH 9.0, 1.5 mM MgCl₂, 0.1% Triton X-100, 0.01% gelatin, 0.2 mM each deoxynucleotide triphosphate and 2.5 units of <u>Thermus aquaticus</u> (Taq) DNA polymerase (Promega Corporation, Madison, WI). The reaction mixture was overlaid with mineral oil. After five cycles (30 seconds at 94°C, 30 seconds at 45°C, 1 minute at 50°C) and twenty-five cycles (30 seconds at 94°C, 30 seconds at 45°C, 1 minute at 72°C) the amplified DNA was removed for analysis.

10

15

20

#### Table 4

## Degenerate Oligonucleotide Primer Sequences (5' - 3')

ZC4519

TTT ATT AGA AAT GTT CTC GGT

ZC4520

CCT CTT CCA TAT TTT TCC ATT

ZC4559

ATA AGA ATT CAT NKR YTT NGC YTC RTT RAA

ZC4560

ATA AGA ATT CTT YRA YGA RAA NGG NGA YGC

ZC4561

ATA AGA ATT CGC NGG NAT HTT YYT NKG NTA

ZC4562

ATA AGA ATT CTA NCM NAR RAA DAT NCC NGC

ZC4563

ATA AGA AAT CAN GTN GTR TAC ATN GTR AA

An aliquot from each reaction was electrophoresed on agarose and transferred to nitrocellulose for Southern analysis. Southern analysis of the PCR products showed that a 460 bp fragment corresponding to the 3' end of the 2b sequence was present in several pools. One of the pools that produced the correct size PCR product encoding the 3' sequence of the 2b subtype was diluted and scr ened with radiolabeled ZC4519 and ZC4520 (Table 4). Phage that hybridize to both radiolab led ZC4519 and ZC4520 are pick d, eluted, diluted, plated and rescreened with the oligonucleotide prob s. The screening is

30

25

35

10

15

20

25

30

35

repeated until a pure clone is obtained. The pure clone is sequenced, and a full-length clone is constructed using the most convenient restriction enzyme(s).

Based on an alignment of the deduced amino acid sequences of subtypes 1a and 2a, strategies were designed for cloning additional subtypes using PCR amplification. Degenerate oligonucleotide families were prepared to encode conserved amino acid sequences in the sixth transmembrane domain, a region surrounding the conserved amino acid sequence Phe-Asp-Glu-Lys, the third cytoplasmic loop, and the second transmembrane domain (Table 4).

Glutamate receptor cDNA sequences were amplified with pairs of degenerate primers from Table 4 using the PCR method on cDNA from the total rat brain library, the cDNA from the rat cerebellum library, a rat cortex cDNA library or a rat hippocampus cDNA library (both obtained from Michael Brownstein, National Institutes of Health, Bethesda, MD). The primers also each contained a 5' tail of 10 nucleotides, which provided convenient restriction enzyme sites. For each PCR reaction, ten nanograms from the library and 100 pmol of the oligonucleotide pools ZC4563 and ZC4560 (Table 4) were combined in a reaction volume of 50 μl containing 50 mM KCl, 10 mM Tris-HCl pH 9.0, 1.5 mM MgCl2, 0.1% Triton X-100, 0.01% gelatin, 0.2 mM each deoxynucleotide triphosphate and 2.5 units of Tag DNA polymerase. The reaction mixture was overlaid with mineral oil. After five cycles (30 seconds at 94°C, 30 seconds at 45°C, 1 minute at 50°C) and twenty-five cycles (30 seconds at 94°C, 30 seconds at 45°C, 1 minute at 72°C) the amplified DNA was removed for analysis.

An aliquot from each reaction was electrophoresed on an agarose gel. Southern analysis of the gel was performed using essentially the method described by Sambrook et al. (ibid.) and random-primed fragments cov ring the entir coding regions from both the subtype la and 2a clones. The autoradiographs showed that the PCR reaction generated fragments of novel size that wer

10

15

20.

25

30

35

diff rent from either the la or 2a subtyp . PCR-generated fragments w r electrophoresed n an agarose gel. Regions corresponding to the unique-sized receptor-related products were excised and electrophoresed onto NA45 paper (Schleicher and Schuell, Keene, NH). The purified fragments were recovered using essentially the method described by the manufacturer, digested with Eco RI and ligated to plasmid pVEGT' that had been linearized by digestion with Eco RI and treated with phosphatase to prevent recircularization. ligation mixtures were transformed into E. coli strain DH10b cells. Transformants were picked and replica plated onto nitrocellulose filters and screened using random-primed probes from the la and the 2a clones. Forty-eight colonies were picked for restriction analysis and sequencing.

DNA sequences from the cDNA from the total rat brain library and the cDNA from the rat cerebellum library were each amplified and analyzed using the methods described above and oligonucleotide ZC4559 in combination with either ZC4561 or ZC4559 (Table 4).

A rat cortex cDNA library and a rat hippocampus cDNA library (both obtained from Michael Brownstein, NIH) are subdivided into 30 pools of 10,000 colonies. Plasmid DNA is prepared from each pool, and the DNA is subjected to Southern analysis after restriction digestion of the pools with Bam HI and Xho I or by PCR amplification of each pool using the degenerate oligonucleotides of Table 4. The library pools containing DNA that hybridize to the probes and appear to contain a full-length cDNA are subdivided. The plasmid DNA is prepared and screened as described above. Positive pools are again divided and the procedure is continued until the pool is reduced to pure clones. The clones are subjected to restriction analysis and partial sequence analysis. Clon s that represent distinct glutamate receptor homologs are complet ly s qu nced. Full length clon s are g n rated by subj cting the original pools to PCR amplification

using an oligonucleotide primer specific to the SP6 promoter at the 5' nd of the cDNA insert and an antisense oligonucleotide primer corresponding to the 5' end of the most complete cDNA to identify pools that contain the longest glutamate receptor homolog cDNA. The pool is then diluted and rehybridized with the probes as described above to isolate a full length cDNA clone.

### Expression of Glutamate Receptor Subtypes

10

15

20

5

Complementary DNA sequences encoding subtypes 1b and 2a were subcloned first into the mammalian expression vector Zem228R to obtain convenient terminal restriction The cDNAs were then subcloned into pVEGT'. The cDNA sequence encoding subtype 1b was constructed by replacing the 3' terminal portion of subtype la described in Example I with the analogous portion of subtype 1b from SN23. Plasmid SN23 was digested with Kpn I and Xba I to isolate the fragment containing the 3' terminus of the 1b subtype. The plasmid containing the subtype 1a coding sequence (45-A) in Zem228R was digested with Kpn I and Xba I to isolate the vector containing fragment. The vector containing fragment is ligated to the Kpn I-Xba I fragment from SN23. The resulting plasmid comprises the MT-1 promoter, the subtype 1b cDNA and the hGH terminator. This plasmid was transfected into the BHK 570 cell line essentially as described in Example I to obtain stably transfected cell lines expressing the subtype 1b receptor. The subtype 1b cDNA fragment was isolated as a Bam HI fragment, which was ligated with pVEGT' that had been linearized with Bam HI. A plasmid containing the cDNA sequence in the correct orientation was used to synthesize RNA in an in vitro system. RNA was injected into occytes as described above.

30

35

25

Plasmid SN30, which comprises the subtype 2a cDNA, was digest d with Eco RI to isolat the subtype 2a cDNA. Th Eco RI fragm nt was ligated with Eco RI-lin arized Zem228R. A plasmid c ntaining the insert in th correct ori ntation was digested with Bam HI to isolate the cDNA

sequence. The Bam HI fragment comprising the subtype 2a cDNA was ligated with Eco RI-linearized pVEGT'. A plasmid containing the cDNA in the correct orientation was used to synthesize RNA in an in vitro translation. The RNA was injected into frog oocytes as described above.

#### EXAMPLE III

## Generation of antibodies to glutamate receptor subtypes

10

15

Receptor subtype-specific polyclonal antisera were generated in rabbits using standard immunization techniques. Synthetic peptides (Table 5) were designed from the cloned receptor sequences. The peptides were conjugated to keyhole limpet hemocyanin, and each antigen was used to immunize two animals. For each peptide, the animals were injected with  $100-200~\mu g$  of conjugated peptide divided among three subcutaneous sites. The animals were immunized at three-week intervals and bled via an ear vein 10 days after the third and subsequent

20

immunizations.

#### Table 5

	,			The state of the s	
05		Subtype	Seq. ID	Peptide Sequence	Apparent Location
25		1a	21	RDSLISIRDEKDGLNRC	extracellular
	. 14	• •	22	DRLLRKLRERLPKARV	extracellular
			23	EEVWFDEKGDAPGRYD	extracellular
Our .			24	EFVYEREGNTEEDEL	cytoplasmic
30			25	PERKCCEIREQYGIQRV	extracellular
	*		26	IGPGSSSVAIQVQNLL	extracellular
		•	27	IAYSATSIDLSDKTL	extracellular
		1b	28	KKPGAGNAKKRQPEFS	cytoplasmic
•	•	FF FF	29	PEFSPSSQCPSAHAQL	cytoplasmic -
.35	. •	2a	30	DKIIKRLLETSNARG	extracellular
		. *	31	VNFSGIAGNPVTFNEN	extracellular
			32	GEAKSELCENLETPAL	cytoplasmic
		2b	33	PARLALPANDTEFSAWV	cytoplasmic
	•				

Anti-peptide antibodies were purified by affinity purification using the ProtonTM Kit (Multiple Peptide Systems (San Diego, CA). Purified antibodies were stored in column elution buffer and neutralizing buffer (supplied by Multiple Peptide Systems). Bovine serum albumin was added to a concentration of 1 mg/ml, and sodium azide was added to a concentration of 0.05%. The antibodies were stored at 4°C or in small aliquots at -20°C.

10

Antibodies generated from the peptides listed in Table 6 were used to detect G protein-coupled glutamate receptors by Western blot analysis of membranes prepared from transfected cell lines that were stably expressing the subtype 1a or subtype 1b receptors. Control cell lines were transfected with vector alone.

15

Table 6
Analysis of Antibodies Raised to Peptides

20	Antibodies to Peptide Sequence	Seq. ID	Location	Western
*	RDSLISIRDEKDGLNRC	21	extracellular	+++ with bkgd
	DRLLRKLRERLPKARV	22	extracellular	+
25	EEVWFDEKGDAPGRYD	23	extracellular	++++ low bkgd
	EFVYEREGNTEEDEL	24	cytoplasmic	++++ low bkgd
	KKPGAGNAKKRQPEFS	28	cytoplasmic	+ for la
*			*.	- for 1b
· · · · · · · · · · · · · · · · · · ·	PEFSPSSQCPSAHAQL	29	cytoplasmic	+++ for 1b low bkgd

30

35

40

Transfectants that were stably expressing either the la or 1b subtype were each grown to confluency in five to ten 150 mm plates. Each plate was first washed twice with 15 ml of cold PBS and then 20 ml of ice cold 10 mM NaHCO3 was added to each plate. The cells from each plate were scraped off th plat s with a rubber spatula and transferred to a glass dounce homogeniz r on ice. The cells w r disrupted with ten strok s of the B pestle. The homog nates from each plat w re c mbin d

10

15

20

25

30

35

and centrifuged for thirty minutes at 3000 rpm at 4°C. The pellets were resuspended in 4-8 ml of 10 mM NaHCO, using a 22 g needle and syringe, and 69% sucrose was added (6-12 ml) to each sample until an index of refraction of 1.410 was reached. The samples were transferred to a high speed centrifugation tube, and each sample was overlayed with 42% sucrose. The samples were centrifuged for two hours at 25,000 rpm at 4°C. The samples were collected by gently floating the membranes off the 42% sucrose layer by adding 1 ml of 10 mM NaHCO, and resuspending the membranes by carefully stirring the The upper layer was transferred to a fresh tube on ice. The purified membranes were centrifuged at 10,000 rpm at 4°C and the pellets resuspended in 10 mM NaHCO. The purified membranes were then adjusted to a final protein concentration of 1-2 µg/ml.

Ten to twenty micrograms of each purified membrane preparations were diluted with 2x SDS-mercaptoethanol buffer (100 mM Tris HCl (pH 6.8), 200 mM dithiothreitol, 4% SDS, 0.2% bromophenol blue, 20% glycerol). samples were incubated for 15 minutes at 37°C followed by boiling for 5 minutes. The samples were subjected to SDS-PAGE on 4-15% gradient gel. The samples were electrotransferred to nitrocellulose using the method essentially described by Towbin (Proc. Natl. Acad. Sci. USA 76: 4350-4354, 1979; which is incorporated herein by reference in its entirety). After transfer, the nitrocellulose was cut into strips such that each strip contained a control and receptor samples. nitrocellulose was preincubated in blocking buffer and then incubated with a dilution of either the preimmune serum or the serum collected after antigenic stimulation (serum from later bleeds (i.e. those after four antigen stimulations) were diluted 1:1500). After washing, a horse radish peroxidase-conjugated goat anti-rabbit antibody (Bi -Rad Laboratories, Richm nd, CA) diluted 1:2,500 was add d and after incubation and washing, the horse radish peroxidase substrate (Bio-Rad Laboratories)

was add d and the color reaction was initiated. The reaction was stopped by rinsing the filters in distilled water. Table 6 shows the results of the Western blot analysis.

Ť.,

67

#### SEQUENCE LISTING

#### (1) GENERAL INFORMATION:

- (i) APPLICANT: Mulvihill, Eileen R. Hagen, Frederick S. Houamed, Khaled M. Almers, Wolfhard
- (ii) TITLE OF INVENTION: G PROTEIN-COUPLED GLUTAMATE RECEPTORS
- (iii) NUMBER OF SEQUENCES: 33
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Townsend and Townsend
  - (B) STREET: One Market Plaza, Steuart Street Tower
  - (C) CITY: San Francisco
  - (D) STATE: California
  - (E) COUNTRY: USA
  - (F) ZIP: 94105-1492
  - (V) COMPUTER READABLE FORM:
    - (A) MEDIUM TYPE: Floppy disk
    - (B) COMPUTER: IBM PC compatible
    - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
    - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 07/672,007
  - (B) FILING DATE: 18-MAR-1991
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 07/648,481
  - (B) FILING DATE: 30-JAN-1991
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 07/626,806
  - (B) FILING DATE: 12-DEC-1990
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Parmelee, Steven W.
  - (B) REGISTRATION NUMBER: 31,990 (C) REFERENCE/DOCKET NUMBER: 13952-6PC
  - (ix) TELECOMMUNICATION INFORMATION:
    - (A) TELEPHONE: (206) 467-9600
    - (B) TELEFAX: (415) 543-5043
- (2) INFORMATION FOR SEQ ID NO:1:

	•	0 .				· .										¥ .		
	•			95					100			11		105			**	*
	TCC Ser	TGC CVs	TGG	CAC	TCT Ser	TCA Ser	Agt	<b>VTG</b>	CTC Leu	GAA Glu	CAG Gln	AGC Ser		GAA Glu	TTC Phe	ATC Ile		745
•			110					115		•		•						
	AGA	GAC	TCC	CTG	ATT	TCC Ser	ATC	CGA	GAT	GAG Glu	AAG Lys	GAT Asp	GGG Gly	CTG Leu	AAC Asn	CGA Arg	*	793
	Arg	125	24t	Dea	116	,	130			٠.	*	135	-		-			• ;
-	maa	CERC .	COTT	САТ	ccc	CAG	ACC	CTG	CCC	CCT	GGC	AGG	ACT	AAG	AAG	CCT		841
	Cys	Leu	Pro	Asp	Gly	Gln	Thr	Leu	Pro	Pro	GIY	Arg	Thr	Lys	Lys	Pro 155	ν.	
	140	• •		. •	,	145		11			150	<i>"</i> ,						
	ATT	GCT	GGA	GTG	ATC	GGC	CCT	GGC	TCC	AGC	TCT	GTG	GCC	ATT	CAA	GTC		889
	Ile	Ala	Gly	Val	Ile 160	GIĀ	Pro	Gly	Ser	Ser 165	Ser	Val	ATA	TTR	170	, AGT		
	CAC	୍ତ ; ଅଧ୍ୟ	ىلىملى	СТС	CAG	CTG	TTC	GAC	ATC	CCA	CAG	ATC	GCC	TAT	TCT	GCC	<u>.</u> .	937
	Gln	Asn	Leu	Leu 175	Gln	Leu	Phe	Asp	Ile 180	Pro	Gln	Ile	Ala	Tyr 185	ser	YIZ		
				- C2 C	- CMC	AGT	GAC	222	ACT	TTG	TAC	AAA	TAC	TTC	CTG	AGG		985
	ACA	Ser	Ile 190	Asp	Leu	Ser	Asp	Lys 195	TIIT	Leu	Tyr	Lys	Tyr 200	Phe	Leu	Arg		
					CAC	ACT	anaric Tarrica	CAG	GCA	AGG	GCG	ĀTG	CTC	GAC	ATA	GTC		1033
	GTG Val	GIC Val	Pro	Ser	Asp	Thr	Leu	Gln	Ala	Arg	Ala	1100	Leu	Asp	Ile	Val	,	
		205			. ".	•	210		•			215				,		
	AAG	· · CGT	י יישר	- AAC	TGG	ACC	TAT	GTC	TCA	GCA	GTC	CAC	ACA	GAA	GGG	AAT Asn	• •	1081
	Lys 220	Arg	Туз	Asr	Trp	Thr 225	Tyr	Val	Ser	Ala	Val 230	****	Thr	Glu	GTÅ	Asn 235		
						እጥር	GAT	GCT	TTC	AAA	GAA	CTG	GCT	GCC	CAG	GAA Glu		1129
	TAC	Gly	GAC Glu	S AG	Gly 240	Met	Asp	Ala	Phe	Lys 245	GTU	Leu	Ala	Ala	Gln 250		47-	
•		•	_ * <u></u>			- 100 A	. m <i>cc</i>	CAC		ልሞሮ	ጥእር	: AGC	AAT	GCT	GGC	GAG Glu		1177
	GGC	CTO	1 Cy	C ATO S Ilo 25!	B Ala	His	Set	Asp	Lys 260	TTE	Tyr	Ser	Asn	Ala 265	Gly	Glu		
٠.	- 6		•						n. ' a a a	· ~~	· ccc	GAG	ccc	CTI	CCC	: AAG		1225
	AAC Lys	AGG Se	r Ph	e Asi	C CGC	Lev	Let	CG1 Arc 275	The	Lev	Arg	Glu	280	,	Pro	AAG Lys		*
	·	*	27			•	•	_						. ccc	: 'GG(	TTA		1273
-	GCC	CAG	G GT	T GT	G GT	C TGC	TT	C TG(	C GAG	GGC	AT(	, ACA t Thi	· Val	. CGG	Gly	TTA Leu		
	Ala	28	g Va 5	. Va.	T Ag	r chs	29	0			, ,,,,,	295	5		`	• • •		
		a - 1 a	m cc	~ <u>a</u> m	e ce	c cad	CT	G GG(	C GT	GT	GGG	C GAC	TT	C TC	CT	C ATT	• •-	1321
	CIV Let	ų. Se	r Al	a Me	t Ar	g Arg	1 re	u Gl	y Val	l Va	Gl;	,	ı Pho	e Sei	r Lei	Ile 315	• .	*
		•	<b>.</b>	m ~~	3 m~	G GC1	V (2)	C AG	A GA'	r GA	A GT	CAT	C GA	A GG	C TA	T GAG T Glu	•	1369
	GG; Gl;	A AG	T GA T As	r GG p Gl	y Tr 32	D. VT	a As	p Ar	g As	p G1:	<u> </u>	1 11	e Gl	u Gl	33	r Glu O	·	
										, -						•	-	

										, •						•		
	Val	Glu	Ala	Asn 335	GŢĀ	GTÄ	TTG	TIIT	ATA Ile 340	<i>-</i> 12		:		345			141	7
	Arg	Ser	Phe 350	Asp	Asp	Tyr	Pne	355	AAG Lys			·	360	:		. "	146	
	Arg	Asn 365	Pro	Trp	Pne	PLO	370	FIIG	TGG Trp	<b>J</b>	••	375			*.	1	151	.3
	CTA Leu 380	CCT	GGA Gly	CAC His	CTC	TTG Leu 385	GAA Glu	AAC Asn	CCC	AAC Asn	TTT Phe 390	AAG Lys	AAA Lys	GTG Val	TGC Cys	ACA Thr 395		
	GGA Fly	AAT I In	GAA Glu	AGC Ser	TTG Leu 400	GAA Glu	GAA Glu	AAC Asn	TAT	GTC Val 405	CAG Gln	GAC Asp	AGC Ser	AAA Lys	ATG Met 410	GGA Gly	160	)9
•	TTT Phe	GTC Val	ATC Ile	AAT Asn 415	GCC Ala	ATC Ile	TAT Tyr	GCC Ala	ATG Met 420	GCA Ala	CAT His	GGG Gly	CTG Leu	CAG Gln 425	AAC Asn	ATG Met	165	57 ,
0	His	His	Ala 430	Leu	Cys	Pro	GIA	435	AGT	GLY	,		440	. *		AAA Lys	· · · ·	
	Pro	Ile 445	Asp	Gly	Arg	ГÄЗ	450	" . Ter	yah		:	455	-•			Phe	17	53
	GTC Val	Gly	GTG Val	TCT Ser	GGA Gly	GAG Glu 465	GAG Glu	GTG Val	TGG		GAT Asp 470	GAG Glu	AAG Lys	GGG	GAT	GCT Ala 475	18	01
	CCC	GGA Gly	AGG	TAI	GAC Asp 480	TTE	ATG Met	AAT Asn	CTG Leu	CAG Gln 485	TAC Tyr	ACA Thr	GAA Glu	GCI Ala	AAT ASI 490	r CGC		49
	TAT	GAC Asp	TAI Tyr	GTC Val	. Hls	GTG Val	GGG	ACC	TGG Trp 500	****	GAA Glu	GGA Gly	GTG Val	Lev 505	AA! Asi	n Ile	18	97
	GAT Asp	GAT Asp	TAC TYI	Lys	ATC	CAG Gln	ATG	AAC Asr 515		AGC Ser	GGA Gly	ATC Met	GTA Val 520	CGA Arg	TC: Sei	r GTG r Val	19	145
	TGC	: AG1 : Se1 : 525	GAC		TGC Cys	TTA Lev	AAG Lys 530	. 47	r cag / Glm	ATI	Lys	GT( Val 535	ATA L-Ile 5	CGC Arg	AA Ly	A GGA B Gly		93
•	GAA Glu	A GTO	•	C TGG	TGC S Cyt	TGG Trp 545	, 114	TGC CY	C ACG	GCC Ala	TGC Cys 550	-	A GAG	3 AAS	r GA n Gl	G TTT u Phe 555		041
٠	GIG	CAC	GA(	C GA	G TT	C ACC	TG	C' AG	A GCC	TG	r GAC	CT	G GG(	g TG	G TG	G CCC	20	089
							SU	BS"	ritl	ITE	SH	EE				·		

71.

٠										٠.									
	Val	Gln	Asp	Glu	Phe 560	Thr	Cys	Arg	Ala	Cys 565	Asp	Leu	Gly	Trp	Trp 570	Pro			
	AAC	GCA	GAG	CTC	ACA	GGC	TGT.	GAG	CCC	ATT	CCT	GTC Val	CGT	TAT Tyr 585	CTT Leu	GAG Glu		2137	
		·		575	•			. ')	200		*		•				· · · · ·		
	TGG	AGT	GÁC	ATA	GAA	TCT	ATC	ATA	GCC	ATC	GCC	TTT	TCT	TGC	CTG	GGC		2185	
	Trp	Ser	Asp 590	Ile	Glu	Ser	Ile	595	ATG	116	WIG	FILE	600						
	ATC	CTC	GTG	ACG	CTG	TTT	GTC	ACC	CTC	ATC	TTC	GTT	CTG	TAC	CGG	GAC	*	2233	
•	Ile	Leu 605	Val	Thr	Leu	Phe	Val 610	Thr	Leu	Ile	Phe	015	Leu	Tyr	Arg	Asp	*		
	A CA	ccc	GTG	GTC	AAA	TCC	TCC	AGT	AGG	GAG	CTC	TGÇ	TAT	ATC	ATT	CTG	•	2281	
•	Thr	Pro	Val	Vál	Lys	Ser 625	Ser	Ser	Arg	Glu	Leu 630	Сув	Tyr	Ile	Ile	Leu 635		•	
	com	CCT	א נוזימו	سابات :	СПС	GGC	TAT	GTG	TGC	CCT	TTC	ACC	CTC	ATC	GCC	AAA		2329	
	Ala	Gly	Ile	Phe	Leu 640	Gly	Tyr	Val	cys	Pro 645	Phe	Thr	Leu	Ile	Ala 650	-1-			
. •			100		mcc.	TCC	TAC	CTC	CAG	CGC	CTC	CTA	GTT	GGC	CTC	TCT	•	2377	
	Pro	ACT	Thr	Thr 655	Ser	Cys	Tyr	Leu	Gln 660	Mrg	Leu	Leu	Val	Gly 665	Leu	Ser	•		
			ame	TICC	ሞልሮ	شاش	<b>ੰ</b> ਫਟਾਾ	тта	GTG	ACC	AAA	ACC	AAT	CGT	ATT	GCA Ala		2425	
	Ser	Ala	Met 670	Cys	Tyr	Ser	Ala	Leu 675	Val	Thr	Lys	Thr	Asn 680	nr 9	Ile	Ala		*	,
			· CTC	COT	GGC	AGC	AAG	AAG	AAG	ATC	TGC	ACC	CGG	AAG	CCC	AGA	•	2473	
	Arg	Ile 685	Leu	Ala	Gly	Ser	Lys 690	ГÄВ	Lys	Ile	Cys	Thr 695	ur A	Lys	Pro	Arg			
	mmo	አጥሮ	acc	COT	TGG	GCC	CAA	GTG	ATC	ATA	GCC	TCC	ATT	CTG	ATT	AGT	. ,	2521	
	Phe 700	Met	Ser	Ala	Trp	Ala 705	GID	Val	Ile	Ile	710	SEL	· Ile	Leu	Ile	Ser 715	•		
	com a		- CTDA	202	. СТА	CTC	стс	ACC	TTG	ATC	ATC	ATG	GAG	CCT	CCC	ATG Met		2569	
	Val	Gln	Leu	The	. Leu	AT	Val	Thr	Leu	Ile   725	, 114	Met	Glu	Pro	Pro 730		*	-0.	
•				•	720	,			٠	7,23				. 8			•	0639	
	CCC	TTA :	TTG	TCC	TAC	CCG	AGT	ATC	AAG	GAA	GTC	TAC	CTI	ATC	TGC	AAT	·, ·	2617	•
	Pro	Ile	Lev	Sex 735	Tyr	Pro	) Ser	. Ile	740	)	. vai	y.		745			•		
٠	ACC	AGO	: AAC	CTG	GGT	GTA	GTG	GCC	CCI	GTG	GG1	TAC	TAA :	GGA	CTC	CTC	0.	2665	
	Thi	Set	750	Leu	Gly	Val	. Val	755	PIC	Va]	Gly	TYI	. Asr 760		, Ter	ı Leu	·		
	አጥ/	ን አጥሶ	· ACC	ጥርባ	ACC	TAC	TAT	GCC	TTC	: AAC	ACC	CGC	AAC	GT	CC	GCC Ala	,	2713	
	Ile	Met 765	: Sez	Cys	Thi	Туг	770	: Ale	Phe	Lys	Thi	775	,	n Val	L PY	Ala	*	300	
							ma.	ን አጥ/	, GC(		c ac	C ATO	TAC	C ACT	C AC	C TGC		2761	
	AA( Asi	n Pho	CAA.S BAB	GAC	Ala	Lys	Ty	Il	Ala	Ph.	Th	r Met	t Ty	r Thi	r Thi	r Cys	سنر	*	

	1		,		785					790					795	•	
780	-				703	٠,				,,,,,					,,,	•	
ATC	ATC	TGG	CTG	GCI	TTC	GTI	CCC	ATT	TAC	TTI	GGG	AGC	AAC	TAC	AAG		2809
Ile	: Ile	Tr	Leu			Val	Pro	Ile	Tyr 805		r GTA	Ser	ASI	810	Lys		
	•			800					903		•	•		. 010	).		
ATC	: ATC	ACI	ACC	TGC	TTC	GCG	GTG	AGC	CTC	AGT	GTG	ACG	GTG	GCC	CTG		2857
Ile	Ile	Thi	Thr	Cys	Phe	Ala	Val	Ser	Leu	Ser	· Val	Thr	Val	Ala	Leu	•	-
		•	815	i		*		820	i- '	•			825			•	
ccc	TGC	àጥር	. Therefore	ACT	CCG	AAG	ATG	TAC	ATC	ATC	ATT	GCC	AAA	CCT	GAG		2905
Gly	Cys	Met	Phe	Thr	Pro	Lys	Met	Tyr	Ile	Ile	Ile	Ala	Lys	Pro	Glu		
_		830		•	-		83,5					840				* .	
300	3: 3.C	CTIC	رخد	ነ አርጥ	GCC	TTC	ACG	ACC	TCT	GAT	GTT	GTC	CGC	ATG	CAC		2953
Ara	Asn	Val	Arg	Ser	Ala	Phe	Thr	Thr	Ser	Asp	Val	Val	Arg	Met	His		
	845					850					855	٠.	•		•		
3	-	.014			CTIC	ccc	ТСС	CGC	TCC	220	ACC	TTC	CTC	220	y draft		3001
GIC Val	GGT	Asn	GIV	Lvs	Leu	Pro	Cys	Arq	Ser	Asn	Thr	Phe	Leu	Asn	Ile		3001
860	_				865			•		870					875		-22-
						000		: 	3.3 m	ccc	3.30	mom.	330	cco	220		2040
TTC	CGG	AGA	AAG	AAG	Pro	GGG	Ala	Glv	AAT	Ala	Asn	TCT	Asn	Glv	Lys	* .	3049
LIIE	wr.A	wi	ت ورد	880		3		,	885					890			
		~4	4	·	*												
TCT	GTG	TCA	TGG	TCT	GAA	CCA	GGT	GGA	AGA	CAG	GCG Ala	CCC	AAG T.ve	GGA	CAG		3097
ser	VAI	Ser	895	261	GIG	PLU	GIJ	900	<b></b> 4	92			905		<b>J</b>		
							٠							. 101	<u> </u>		
CAC	GTG	TGG	CAG	CGC	CTC	TCT	GTG	CAC	GIG	AAG	ACC	AAC	GAG	ACG	GCC		3145
H1S	Val	910		Arg	reu	Sel	915	nis	, AT	ri A a	. 1111	Asn 920	GIU	1111	VIG		•
						•		•					<u> </u>			•	
TGT	AAC	CAA	ACA	GCC	GTA	ATC	AAA	CCC	CTC	ACT	AAA	AGT	TAC	CAA	GGC		3193
Cys		Gln	Thr	ATA	VAI	330	rys	PLO	Tea	TILL	935	Ser	TAT	GIII	GIY		0.
	925				-				•					÷ .			
TCT	GGC	AAG	AGC	CTG	ACC	TTT	TCA	GAT	GCC	AGC	ACC	AAG	ACC	CTT	TAC	٠.	3241
	Gly	Lys			Thr 945	Pne	Ser	Asp	ATA	950	Inr	Lys	Thr	reu	955	;	
940			. 4			0				•				•		. , ,	
AAT	GTG	GAA	GAA	GAG	GAC	AAT	ACC	CCT	TCT	GCT	CAC	TTC	AGC	CCT	CCC		3289
Asn	Val	Glu	Glu		Asp	Asn	Thr	Pro	Ser	Ala,	His	Phe	Ser	PTO 970	Pro		
÷ .				960					700					3/0	•		
AGC	AGC	CCT	TCT	ATG	GTG	GTG	CAC	CGA	CGC	GGG	CCA	CCC	GTG	GCC	ACC		3337
Ser	Ser	Pro	Ser	Met	Val	Val	His	Arg	Arg	Gly	PTO	Pro	Val	Ala	Thr		
			975					980				•	985				*
ACA	CCA	CCT	CTG	CCA	CCC	CAT	CTG	ACC	GCA	GAA	GAG	ACC	CCC	CTG	TTC		3385
Thr	Pro	Pro	Leu	Pr	Pro	His	Leu	Thr	Ala	Glu	Glu	Thr	Pro	Leu	Phe		-
- <del></del>	=	990				•	995				,	1000	)	•	-		
~~~	c cm	Cam	ጥርር	CTC	יאיים	CCC	AAG	GGC	TTG	CCT	CCT	CCT	CTC	CCG	CAG	•	3433
Lan	Ala	YED	Ser	Val	Ile	Pro	Lys	Gly	Leu	Pr	Pr	Pr	Leu	Pro	Gln		
	1005		-			1010				•	1015	5			•		·•·

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1199 amino acids

- (B) TYPE: amino acid
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Val Arg Leu Leu Ile Phe Phe Pro Met Ile Phe Leu Glu Met Ser Ile Leu Pro Arg Met Pro Asp Arg Lys Val Leu Leu Ala Gly Ala

Ser Ser Gln Arg Ser Val Ala Arg Met Asp Gly Asp Val Ile Ile Gly

Ala Leu Phe Ser Val His His Gln Pro Pro Ala Glu Lys Val Pro Glu

Arg Lys Cys Gly Glu Ile Arg Glu Gln Tyr Gly Ile Gln Arg Val Glu

Ala Met Phe His Thr Leu Asp Lys Ile Asn Ala Asp Pro Val Leu Leu

Pro Asn Ile Thr Leu Gly Ser Glu Ile Arg Asp Ser Cys Trp His Ser 100

Ser Val Ala Leu Glu Gln Ser Ile Glu Phe Ile Arg Asp Ser Leu Ile

Ser Ile Arg Asp Glu Lys Asp Gly Leu Asn Arg Cys Leu Pro Asp Gly

Gln Thr Leu Pro Pro Gly Arg Thr Lys Lys Pro Ile Ala Gly Val Ile 150 145

Gly Pro Gly Ser Ser Ser Val Ala Ile Gln Val Gln Asn Leu Leu Gln

Leu Phe Asp Ile Pro Gln Ile Ala Tyr Ser Ala Thr Ser Ile Asp Leu 180

Ser Asp Lys Thr Leu Tyr Lys Tyr Phe Leu Arg Val Val Pro Ser Asp 195

Thr Leu Gln Ala Arg Ala Met Leu Asp Il Val Lys Arg Tyr Asn Trp 210

Thr Tyr Val S r Ala Val His Thr Glu Gly Asn Tyr Gly Glu Ser Gly 230

				•					• 00	, ,			19.			
į.					245					250						
Н	is	Ser	Asp	Lys 260	Ile	Tyr	Ser	Asn	Ala 265	Gly	Glu	Lys	Ser	Phe 2	Asp A	Arg
ı	eu	Leu	Arg 275	Lys	Leu	Arg	Glu	Arg 280	Leu	Pro	Lys	Ala	Arg 285	Val '	Val '	Val
C	:ys	Phe 290	Cys	Glu	Gly	Met	Thr 295	Val	Arg	Gly	Leu	Leu 300	Ser	Ala	Met .	Arg
	irg 305		Gly	Val	Val	Gly 310	Glu	Phe	Ser	Leu	Ile 315	Gly	Ser	Asp	Gly	Trp 320
2	lla	Asp	Arg	Asp	Glu 325	Val	Ile	Ğlu	Gly	Tyr 330	Glu	Val	Glu	Ala	Asn 335	Gly
		•	•	_. 340		¥			345					Phe 350		*
			355					200	• •					Pro		
•		370)				3/5	•		÷		300		Gly	•	
	385					390					333	· · ·		Glu		
		•			405	•			٠.	410			141	Ile	-	
				420)		٠.		423	٠,,		•	. *	Ala 430	•	•
		. ;	43	5		•		440	ı		• •		110			
	•	45	0 ,		-• '.		40:	· . · .			•	-,00		Val		
•	465	5			-30-	470) '				7/2		1	Arg		
			•		48	5				450		•		Tyr	• • • •	
				50	0		-,	, ,	. 50.)	•	- '	•	Tyr 510	. :	• -
	Gl	n Me	t As 51	n Ly 5	s Se			221	U -							
	Lei	L Ly 53		y Gl	n Il	Ly	s Va 53		e Ar	g Lys	s Gly	7 Glu 540	ı Val	l Ser	Cys	Cys

Trp	-Ile	Cys	Thr	Ala	Cys	Lys	Glu	Asn	Glu	Phe 555	Val (Gln	Asp. (Glu	Phe 560
545			·		JJU				Trp 570						
				202	•				•		٠. `		1		
•••			580					300	Leu		·=		•		
		595					000		Leu						
,	610	٠			•	0.13	,	*	Arg	•					
625					930	•			Ile	_					
				645			,		Ala 650		•				
			660					000	Leu					•	
		675					000	•			٠.				Gly
	690) ·				093	ı		Pro						
705	5				710	, .									Leu 720
				/23)				Pro 730						
			. 740	,							•		•		Gly
		75	5 3				/ 00	•			,				Thr
	770)				//-	• • •							•	Ala
Ly:	s Ty: 5	r Ile	e Ala	a Phe	Th:	r Met	Туг	Thi	Thr	795	·Ile				Ala 800
Ph	e Va		·	809	5				1 Tyr 810	,				815	
			82	U			•	•							. Thr
Pr	o Ly	s Me 83	t Ty: 5	r Il	e Il	e Il	e Ala 84	Ly:	s Pr	Glu	1 Arg	845	val	. Arg	ser

			1		•					1	•••	<u> </u>	1	<u>ما</u>	Tue
•	850		4	Ser	1	855					000				
865		•	r ii		870					6/3		*			
Pro	Gly	Ala	Gly	Asn 885	Ala	Asn	Ser	Asn	Gly 890	Lys	Ser	Val	Ser	Trp 895	Ser
Glu	Pro	Gly	Gly 900	Arg	Gln	Ala	Pro	Lys 905	Gly	Gln	His	Val	Trp 910	Gln	Arg
Leu	Ser	Val 915	His	Val	Lys	Thr	Asn 920	Glu	Thr	Ala	Cys	Asn 925	Gln	Thr	Ala
Val	Ile 930	Lys	Pro	Leu	Thr	Lys 935	Ser	Tyr	Gln	Gly	Ser 940	Gly	Lys	Ser	Leu
Thr 945	Phe	Ser	Asp	Ala	Ser 950	Thr	Lys	Thr	Leu	Tyr 955	Asn	Val	Glu	Glu	Glu 960
Asp	Asn	Thr	Pro	Ser 965	Ala	His	Phe	Ser	Pro 970	Pro	Ser	Ser	Pro	Ser 975	Met
Val	Val	His	Arg 980	Arg	Gly	Pro	Pro	Val 985	Ala	Thr	Thr	Pro	Pro 990	Leu	Pro
Pro	His	Leu 995		Ala	Glu	Glu	Thr 100	Pro 0	Leu	Phe	Leu	Ala 100	Asp 5	Ser	Val
Ile	Pro		Gly	Leu	Pro	Pro 101	Pro 5	Leu	Pro	Gln	Gln 102	Gln O	Pro	Gln	Gln
Pro		Pro	Gln	Gln	Pro	Pro 0	Gln	Gln	Pro	Lys 103	Ser 5	Leu	Met	Asp	Gln 1040
Leu	. Gln	Gly	Val	Val 104	Thr 5	Asn	Phe	Gly	Ser 105	Gly	Ile	Pro	Asp	Phe 105	His 5
Ala	Val	Leu	Ala 103	. Ģly	Pro	Gly	Thr	Pro 106	Gly	Asn	Ser	Lev	Arg	Ser O	Leu
Tyr	Pro	Pro 107) Pro	Pro	Pro	Gln 108	His O	Lev	ı Glr	n Met	Lev 108	Pro	Leu	His
Lev	Ser 109		- Phe	≆ Gln	Glu	Glu 109	Ser 5	· Ile	Ser	Pro	Pro 110	Gly	y Glu	a Asp	Ile
Asp 110) Asi	Sei	r Glu	Arg 111	Phe O	Lys	Lev	Let	1 Glr 111	ı Glu 15	Phe	e Val	Туг	Glu 1120
		ı Gly	/ Asi	n Thr 112	Glu S	Glu	ı Ast	Glu	1 Let	ı Glu 30	ı Glu	ı Glu	ı Glı	ASY 11:	Leu 35
Pro	Thi	r Ala	s 5 1	r Lys 40	Lev				43			Ala	Let 11	1 Thi	r Pr
•			i .		. •	SU	BSTI	TUT	E SH	EET					

. 78

Pro Ser Pro Phe Arg Asp Ser Val Ala Ser Gly Ser Ser Val Pro Ser

Ser Pro Val Ser Glu Ser Val Leu Cys Thr Pro Pro Asn Val Thr Tyr 1170 1180

Ala Ser Val Ile Leu Arg Asp Tyr Lys Gln Ser Ser Ser Thr Leu 1185 1190 1195

- (2) INFORMATION FOR SEQ ID NO:3:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - 'ii) MOLECULE TYPE: CDNA
 - (vii) IMMEDIATE SOURCE: (B) CLONE: ZC775
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCTAGCATAA CCCCTTGGGG CCTCTAAACG GGTCT

- (2) INFORMATION FOR SEQ ID NO:4:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 43 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: CDNA
 - (vii) IMMEDIATE SOURCE: (B) CLONE: ZC776
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CTCAAGACCC GTTTAGAGGC CCCAAGGGGT TATGCTAGCT GCA

- (2) INFORMATION FOR SEQ ID NO:5:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 40 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: CDNA

SUBSTITUTE SHEET

2 5

. .

(vii)	IMME	DIATE	SC	URCE:
•	(B)	CLONE	:	ZC777

(XI) SEQUENCE DESCRIPTION: SEQ ID NO:5: TGAGGGGTTT TTTGCTGAAA GGAGGAACTA TGCGGCCGCA

40

- (2) INFORMATION FOR SEQ ID NO:6:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 40 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: CDNA
 - (vii) IMMEDIATE SOURCE: (B) CLONE: ZC778
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6: AGCTTGCGGC CGCATAGTTC CTCCTTTCAG CAAAAAACCC

40

- (2) INFORMATION FOR SEQ ID NO:7:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: CDNA
 - (vii) IMMEDIATE SOURCE:
 (B) CLONE: ZC1751
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AATTCTGTGC TCTGTCAAG

3

- (2) INFORMATION FOR SEQ ID NO:8:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: lin ar
 - (ii) MOLECULE TYPE: CDNA

- (vii) IMMEDIATE SOURCE: (B) CLONE: ZC1752

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8: GATCCTTGAC AGAGCACAG

22

- (2) INFORMATION FOR SEQ ID NO:9:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: CDNA
 - (vii) IMMEDIATE SOURCE: (B) CLONE: ZC2063
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GATCCAAACT AGTAAAAGAG CT

- (2) INFORMATION FOR SEQ ID NO:10:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 14 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: CDNA
 - (vii) IMMEDIATE SOURCE: (B) CLONE: ZC2064

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CTTTTACTAG TTTG

- (2) INFORMATION FOR SEQ ID NO:11:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 43 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA

(vii)	IMMEDI	ATE	SOU	JRCE	::	
	(B) C	LONE	: 2	ZC29	3	8

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11: GACAGAGCAC AGATTCACTA GTGAGCTCTT TTTTTTTTT TTT

43

- (2) INFORMATION FOR SEQ ID NO:12:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA
 - (Vii) IMMEDIATE SOURCE: (B) CLONE: ZC3015
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TTCCATGGCA CCGTCAAGGC T

- (2) INFORMATION FOR SEQ ID NO:13:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: CDNA
 - (Vii) IMMEDIATE SOURCE: (B) CLONE: ZC3016
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AGTGATGGCA TGGACTGTGG T

- (2) INFORMATION FOR SEQ ID NO:14:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: CDNA

.

21

: 21

	ĺ		· ·		* *		
(vii)	IMMED:	TATE SOURCE:	. 1				
•	(B) (CLONE: ZC3652					. •
	***		•		* *		
	ή		SEO TI	NO:14:			
(xi)	SEQUE	NCE DESCRIPTION:	SEG II	7 10 1 14 1		v ·	*
		nonamam		-			20
CATGCACC	A TGC	ICIGIGI			•	•	
2) INFOR	OITAM	N FOR SEQ ID NO:	15:	. 1	*		
(4)	CPOTTE	NCE CHARACTERIST	rics:			. You	•
(1)	/A) :	LENGTH: 21 Dase	barra		*	•	
٠	(B)	Type: nucleic ac	310	, -			•
	(C)	STRANDEDNESS: S	rudie				
	(D)	TOPOLOGY: linear			130		. *
	ι-,			•			
(ii)	MOLEC	ULE TYPE: CDNA		•	1		0
(/		7				•	- 2
100	•				(A) (A)		
(vii)	IMMED	IATE SOURCE:	•		•		
,/	(B)	CLONE: ZC3654		•	0	<i>v</i>	
• .				•	r .		
			×		141		•.
(xi)	SEQUE	NCE DESCRIPTION	: SEQ I	D MO:T2:	* 4		÷ (1)
		• • • • • • • • • • • • • • • • • • • •					21
GTGATGG	CA TGG	ACTGTGG T			. · · · · · · · · · · · · · · · · · · ·		•-
	4		.16.				
2) INFO	RMATIC	N FOR SEQ ID NO	: 10:				
	•				*	•	
(1)	SEQUE	NCE CHARACTERIS	ge nair	'S	• .		0
	(A)	LENGTH: 5236 ba	Cid	-	·		
•	(B)	TYPE: nucleic a STRANDEDNESS: S	ingle		*		•
	(°C)	PODOLOGY: 14 POB	7 -~				
	(D)	TOPOLOGY: linea		300			
	****	THE MUDE: ONLY					A."
(11)	MOLE	TULE TYPE: CDNA					-
		,				s en	
•		THE SALIBOR.			- ()		
(vii)	IMME	DIATE SOURCE:	· m		<i>p</i>		
,	(B)	CLONE: SN23	*		• • • • •		1
	•						
(ix)	FEAT	JRE:					
:	(A)	NAME/KEY: CDS	3344			·	• . '
, n.e.	(B)	LOCATION: 627.		•	1.1	•	• •
•					,		
			i. CFA '	TD NO:16		eYe	*
(xi)	SEQU	ENCE DESCRIPTION	i and		•	. 5	
			Alakhili <u>Vila</u>	TGGTGATG	CCTTTGTGTC	TACAGGGCAC	60
AAGAATI	TT - AT	AAATACTC TGGGAAT	TITT WT	-24+413-4	, ·,		, <u></u>
		GCTCTGGT GTGAAGT	יים איי פרי	GGGACTTG	TGGCTAGAGA	AGCTTTTCAA	120
ACGTTCCA	iga ga	GETETGGT GTGAAG	ravi aa		-	·	
		TOTAL TOTAL	ACAC AC	GTCTGAGG	TTCTCAACAT	CAGAGCAGAG	180
rggcctta	LAA CT	CTGGGTCC TGCTTG	HUNG NO		. -		
		TOTAL MOORES	ברכר כר	ACTTOTOA	ACACTTAGTO	CTCTGATCGG	240
		MATERIAL TO A CONTRACTOR AND A CONTRACTO	JULY VV				

AUDITITE SHEET

TGCCTGCGAA CCGAGAACGG CTGCAGTCCT CTGACCTGAG ACCAATAGCT GTGTCTACCC

GGACTCAGCG TCCAGCTCAC CGCCACTAAC GCGCCGCGCA TTGGACACCT GATCCACACA	360
CCTTCGGGCA CCAGTGAAAA ACCGCGACTT GATTTTCTGG AAGAACGCCC CCAGGGTGTG	420
GGAGCGGTCG TGGAGGACCA GCAGGAGGAA GCGGAGGGGA GAGGGGCAGT AGTGGAGGCA	480
GAGAAAGCGT TGAACCAGCT GTGTTGGCCG AAGGCACGAA ACGGCAAAAG GCAGCGGTGA	540
GAGARAGEGT TGARECAGET CICLEGET GAGARETTGE AGGERGACE GGEGTGGGAA CGTGGETGGE GCATCTGTGT GGTTCCCGCT GGGARCCTGC AGGCRGACE GGCGTGGGAA CGTGGCTGGC	600
ATC CTC CTC TTG ATT TTC TTC	653
Met Val Arg Leu Leu Ile Phe Phe 1 5	
CCA ATG ATC TIT TTG GAG ATG TCC ATT TTG CCC AGG ATG CCT GAC AGA	701
Pro Met Ile Phe Leu Glu Met Ser Ile Leu Pro Arg Met Pro Asp Arg 10 20 25	
AAA GTA TTG CTG GCA GGT GCC TCG TCC CAG CGC TCC GTG GCG AGA ATG	749
AAA GTA TTG CTG GCA GGT GCC 1CG 1CC CAS SET Val Ala Arg Met Lys Val Leu Leu Ala Gly Ala Ser Ser Gln Arg Ser Val Ala Arg Met 30 35	
GAC GGA GAT GTC ATC GGA GCC CTC TTC TCA GTC CAT CAC CAG CCT Asp Gly Asp Val Ile Ile Gly Ala Leu Phe Ser Val His His Gln Pro	797
45	045
CCA GCC GAG AAG GTA CCC GAA AGG AAG TGT GGG GAG ATC AGG GAA CAG Pro Ala Glu Lys Val Pro Glu Arg Lys Cys Gly Glu Ile Arg Glu Gln 60 65	845
TAT GGT ATC CAG AGG GTG GAG GCC ATG TTC CAC ACG TTG GAT AAG ATT Tyr Gly Ile Gln Arg Val Glu Ala Met Phe His Thr Leu Asp Lys Ile 75 80 85	893
AAC GCG GAC CCG GTG CTC CTG CCC AAC ATC ACT CTG GGC AGT GAG ATC ASN Ala Asp Pro Val Leu Pro Asn Ile Thr Leu Gly Ser Glu Ile 90 95	941
CGG GAC TCC TGG CAC TCT TCA GTG GCT CTC GAA CAG AGC ATC GAA Arg Asp Ser Cys Trp His Ser Ser Val Ala Leu Glu Gln Ser Ile Glu 2. 110	989
TTC ATC AGA GAC TCC CTG ATT TCC ATC CGA GAT GAG AAG GAT GGG CTG Phe Ile Arg Asp Ser Leu Ile Ser Ile Arg Asp Glu Lys Asp Gly Leu 135	1037
AAC CGA TGC CTG CCT GAT GGC CAG ACC CTG CCC CCT GGC AGG ACT AAG Asn Arg Cys Leu Pro Asp Gly Gln Thr Leu Pro Pro Gly Arg Thr Lys 140	1085
AAG CCT ATT GCT GGA GTG ATC GGC CCT GGC TCC AGC TCT GTG GCC ATT Lys Pro Ile Ala Gly Val Ile Gly Pro Gly Ser S r Ser Val Ala Ile 155 160 165	1133
CAA GTC CAG AAT CTT CTC CAG CTG TTC GAC ATC CCA CAG ATC GCC TAT Gln Val Gln Asn Leu Gln Leu Phe Asp Ile Pr Gln Ile Ala Tyr 180	1181

						•												
TC Se	T (GCC Ala	ACA Thr	AGC Ser	ITE	GAC Asp	CTG . Leu	AGT (.,	AAA Lys 195	ACT Thr	TTG Leu	TAC Tyr	AAA Lys	TAC Tyr 200	TTC Phe	13	229
CI	G .	AGG Arg	GTG Val	Val	bio	TCT Ser	GAC Asp	ACT Thr		C) C	GCA Ala	agg Arg	GCG Ala	ATG Met 215	CTC Leu	GAC Asp	1	277
				205		-			ms/M	CTIC	mCA	GCA	GTC	CAC	ACA		1	325
			220					,	a 3 M	COM	מיניר	222	GAA	CTG	GCT	GCC	1	373
G]	L y	Asn 235	Tyr	Gly	GIR	Ser	240	Mer	moc	Ġ3.C	***	245	тас	AGC	AAT	GCT	1	.421
2	ln 50	clu	Gly	Leu	Cys	255	MIG				260		• •			GCT Ala 265		
G(SC Ly	GAG Glu	AAG Lys	AGC Ser	TTT Phe 270	GAC Asp	CGG Arg	CTC Leu	CTG Leu	CGT Arg 275	AAA Lys	CTC	CGG	GAG Glu	Arg 280	CTT	1	L469
C	CC	AAG Lys	GCC Ala	AGG Arg 285	Agt	GTG Val	GTC Val	TGC Cys	TTC Phe 290	TGC Cys	GAG Glu	GGC Gly	ATG Met	ACA Thr 295	GTG Val	CGG Arg	1	1517
G G	GC ly	TTA Leu	CTG Leu	Ser	GCC	ATG Met	CGC Arg	CGC Arg 305	CTG	GGC	GTC Val	GTG Val	GGC Gly 310	GAG Glu	TTC	TCA Ser	*	1565
C	TC eu	ATT Ile 315	GGA Gly		GAT Asp	GGA Gly	TGG Trp 320	NT.	GAC Asp	AGA Arg	GAT Asp	GAA Glu 325	GTC Val	ATC Ile	GAA Glu	GGC		1613
I	yr	GAG Glu		GAP	GCC Ala	AAC Asn 335	GTA	GGG	ATC Ile	ACA Thr	ATA	AAG Lys	CTI Lev	CAG Gln	TCI	CCA Pro 345		1661
	30 AG	•	AGC	TC! Sel	. Pne	GAT Asp		TAC Tyr	TTC	CTG Lev 355	AAG Lys	CTG Leu	AGG Arg	CTG Lev	GAC Asi 360	ACC Thr	· · · ,	1709
2	LAC LSI	ACI Thi	A AGG	G AAS G ASI 36!	r cci n Pro	TGG Tr	TTC Phe	CCI Pro	GAG G10 370	TTC	TGG Tr	G CAP	A CAT	r CGC s Arg	TTC Phe	CAG Gln).	1757
	Cys	Ar	g Lei 38	u Pro O	o GI	A ur:	5 1000	385	5			٠.	39	0		A GTG s Val		1805
(Cys	Th:	r Gli 5	y as:	U GT	u se.	400		_	_		40	5			C AAA r Lys		1853
	ATC	GG.	A TT	T GT	C AT	C AA	r gc	C ATO	C TA	T GC	C AT	g gc	A CA	T GG	ct S	G CAG	ITUTE	1901 SHEET

Met-	Gly	Phe	Val	Ile	Asn 415	Ala	Ile	Tyr	Ala	Met 420	Ala	His	Gly	Leu	Gln 425	ψ - Σ
AAC Asn	ATG Met	CAC His	CÁT His	GCT Ala 430	CTG Leu	TGT Cys	CCC Pro	GGC Gly	CAT His 435	GTG Val	GGC Gly	CTG Leu	TGT Cys	GAT Asp 440	GCT Ala	1949
ATG Met	AAA Lys	CCC Pro	ATT Ile 445	GAT Asp	GGC Gly	AGG Arg	AAG Lys	CTC Leu 450	CTG Leu	GAT Asp	TTC Phe	CTC Leu	ATC Ile 455	AAA Lys	TCC Ser	1997
TCT Ser	TTT Phe	GTC Val 460	GGA Gly	GTG Val	TCT Ser	Gly	GAG Glu 465	GAG Glu	GTG Val	TGG Trp	TTC Phe	GAT Asp 470	GAG Glu	AAG Lys	GGG Gly	2045
Asp	GCT Ala	CCC	GGA Gly	AGG Arg	TAT Tyr	GAC Asp 480	ATT Ile	ATG Met	AAT Asn	Leu	CAG Gln 485	TAC Tyr	ACA Thr	GAA Glu	GCT Ala	2093
AAT Asn 490	CGC Arg	TAT	GAC Asp	TAT	GTC Val 495	CAC His	GTG Val	GGG Gly	ACC	TGG Trp 500	CAT His	GAA Glu	GGA Gly	GTG Val	CTG Leu 505	2141
AAT Asn	ATT Ile	GAT Asp	GAT Asp	TAC Tyr 510	AAA Lys	ATC Ile	CAG Gln	ATG Met	AAC Asn 515	гÃа	AGC Ser	GGA Gly	ATG Met	GTA Val 520	CGA Arg	2189
TCT Ser	GTG Val	Cys	AGT Ser 525	Glu	CCT Pro	TGC Cys	TTA	AAG Lys 530	GLY	CAG Gln	ATT	AAG Lys	GTC Val 535		CGG	2237
AAA Lys	GGA Gly	GAA Glu 540	Val	AGC Ser	TGC Cys	TGC Cys	TGG Trp 545	TIE	TGC	ACG Thr	GCC Ala	TGC Cys 550	AAA Lys	GAG Glu	AAT Asn	2285
GAG Glu	TTT Phe 555	Val	CAG Gln	GAC Asp	GAG Glu	TTC Phe 560	Thr	TGC Cys	AGA Arg	GCC	TGT Cys 565	vob	CTG	GGG Gly	TGG	2333
TGG Trp 570	CCC Pro		GCA Ala	GAG Glu	CTC Leu 575	Thr	GGC	TGI Cys	GAG Glu	CCC Pro 580	TTE	CCI Pro	GTC Val	CGT	TAT TYP 585	2381
		TGG Trp	AGT Ser	GAC Asp 590	Ile	GAA Glu	TCI	ATC	ATA	3 WIG	ATC	GCC Ala	TTI Phe	Ser 600	TGC Cys	2429
CTG Leu	GGC Gly	ATC	CTC Lev	ı Val	ACC L Thr	CTG Lev	TTI Phe	GTG Val	LTIM	CTC Leu	ATC	TTO Phe	GTI Val		TAC Tyr	2477
CGG Arg	GAC Asp	ACA	ccc Pr		GTC Val	AAA L Lys	TC0 Se1 62!	. Se	C AGT	r Ago r Aro	GAC GAC	CTC Let 630		TAT	T ATC	2525
ATT Ile	CTG		n cc¶	T AT	r TT e Ph	CTC	: GG(TA'	r GT(r Va.	G TGC	CC.	r TTC Ph	C ACC	CTC	C ATC	2573

							640	*			•	645	, , , , , , ,					
,	•	635	٠,		* .				•				OMO	COT N	CUM	GGC	26	21
	GCC	AAA Lys	CCT Pro	ACT	ACC Thr	ACA Thr	TCC Ser	TGC Cys	TAC Tyr	CTC Leu	CAG Gln : 660	Arg	CTC	Leu	Val	Gly 665		- 2.
	650				5 - 59	622				•			. "	_			26	69
	CTIC	ा चा∕्या	TCT	GCC	ATG -	TGC	TAC	TCT	CCT	TTA	GTG .	ACC	AAA Lys	ACC	AAT	CGT	20	903
	Leu	Ser	Ser	Ala	Met	Cys	Tyr	Ser	Ala	Leu 675	Val	TNE	Lys	TIIT	680	nr.		
					6.70												2.	717
	አጥጥ	GCA	CGC	ATC	CTG	GCT	GGC	AGC	AAG	AAG	AAG	ATC	TGC	ACC	CGG	AAG	; 2	717
	Ile	Ala	Arg	Ile 685	Leu	Ala	Gly	Ser	Lys 690	Lys	Lys	ITE	Cys	695	ALY	1 13		
							mcc.	CCC	CAA	GTG	ATC	ATA	GCC Ala	TCC	ATT	CIG	, 21	765
	CCC	AGA	TTC	ATG	AGC	Ala	Tro	Ala	Gln	Val	Ile	Ile	Ala 710	Ser	Ile	Leu		
•	PIO	Arg	700	Mer	261			705	•				710					
							· ·	CTIC	CTC:	ACC	ттG	ATC	ATC	ATG	GAG	CCT	2	813
	ATT	AGT	GTA	CAG	CTA	ACA Thr	Leu	Val	Val	Thr	Leu	Ile	Ile	Met	Glu	Pro		٠.
:	Ile	5er 715	Val	GIU	Ter	144-	720	·		•		725			• .	*	-	
			á					-	3 CT	እጥሮ	DAG	GAA	GTC	TAC	CTT	ATC	; 2	861
	CCC	ATG	CCC	TTA	TTG	TCC	TAC	Pro	Ser	Ile	Lys	Glu	GTC Val	Tyr	Leu	Ile 745		
	Pro 730	Met	PTO	TTG	Feu	735	-1-	•	-	* .	740		•	8		745	ب جنہ ج	
	-		ī		•			cm.	CITIC:	GCC	CCT	GŤG	GGT	TAC	AAT	GGA	2	909
	TGC	AAT	ACC	AGC	AAC	CTG	GGT	Val	Val	Ala	Pro	Val	Gly	Tyr	Asn	GGA Gly		
	Cys	Asn	THE	Ser	750		ē	•,		755					760			
				•				. ms.c	መልጥ	GCC	TTC	AAG	ACC	CGC	AAC	GTG Val	2	957
	CTC	CIC	ATC	ATG	AGC	TGT	Thr	TVI	Tyr	Ala	Phe	Lys	Thr	Arg	Asn	Val		
	Leu	Leu	TTe	765	261	-	, -,		770		•		• • •	775			1	
		• •							ጥልሮ	ATC	GCC	TTC	ACC	AŢĢ	TAC	ACT Thr	3	005
	CCG	GCC	AAC	TTC	AAT	GAG	Ala	LVS	Tyr	Ile	Ala	Phe	Thr	Met	Tyr	Thr	• •	e
	Pro	ALG	780) Pile	, goi	-		785			•		790					
-			•	-			COT		- दक्त	ccc	ATT	TAC	TTT	GGG	AGC	AAC Asn		3053
•	ACC	TGC	ATC	ATC	TGG	CTG	Ala	Phe	Val	Pro	·Ile	Tyr	Phe	Gly	Sei	: Asn		
	Thi	. cys 795	TTE	ت. ۲۳۵	, ,		800		• .*		٠.	805	,	40	. •	•		
								·	CCC	: С ТС	AGC	CT	AGT	GTG	ACC	GTG Val		3101
	TAC	: AAG	ATC	ATC	ACI	ACC	: TGC	Phe	Ala	Val	Ser	Lev	Ser	· Val	Th	7 Val 825		٠.
			; IIe	3 TT#	8 Titr	815	5				820					625		
	810			•		. <u> </u>		- 000	. 220	: አጥር	: ጥልር	ATO	ATC	ATI	r GC	CAAA		3149
	GC	CIC	GG	TGC	ATC	TT	C ACT	r CCC	LVS	Met	Tyr	110	e Ile	Ile	Al	C AAA a Lys O	. '	
	Ala	Lev	1 G1;	y Cys	830) FIIe	3 2000	,		835	5				84	0	٠	
				•			· 		 - mm	 	: <u> </u>	. тс	r gai	GT:	r gt	C CGC		3197
	CC	C GAC	AG	G AAC	GT	CGC	AG	r GCC - Als	- Tit	a Thi	Thi	Se	r Asi	Va.	l Va	c cgc l Arg	`	
-	Pr	Gl	1 Ar	g Asi 84!	U Agr	LAF			850					85	5 .	*		
							_ = -			. ma	n <i>c</i> cc	• ጥር	C AAC	C AC	C TT	C CTC	0	3245
	AT	G CAC	c GT	C GG	r GA	r GG	CAA	R CIY	i CC(u Pr	TG(s Arc	Se	r Asi	n Th	r Ph	c ctc e Leu		
	Me	t His	s Va	T GT	y As	b "GT;	Å πλι	86		٠, ٢	;	-	87	Ó	•		~	•
			86	J							,							

			•			
			88		*	*
racccacca	GCGAGAGGG	CAGGAGGCGG	GGGTAGGTTC	GGACAACAGC	TCCCATCTCA	4834
	CTCCTGAGTC	TTCAGACTCC	TGGACTAAGG	AAGACCCGGG	GACTGACCTT	4894
ACCIIGACI	CHUTCCACTG	CTGTGATCCA	TTGCCAGCCT	GTAGTCACCC	GGGATAAAGG	4954
ATGAGGGTCC	THE CAPTO	СТСТСАТТСС	CTGTGTTTAA	GGAAAAGGAA	AGTATGAGCA	5014
CACAGTAACC	CHARAGEC	CCCATTAGAA	GTTACGGGGG	AGAAAAAAG	AGAAGCAAGA	5074
AAGCTATCAC	CAAAAAAGC	TTGAACAAGG	TGAGCGTGCT	TCACAGATTC	CGTATTAATG	5134
rgatatataa	mmmcc \ C \ CG	AGAAGATAA	CAAGGAGTGT	CAGGCCGTTT	GTGAACTCAC	5194
		•	CCTTCAGCAA	·		5236
TTGCACTGTG	CCAACCAGGI	TCTCCGCTGC			(3)	
(2) INFORM	ATION FOR S	EQ ID NO:17	:	(*	·	
	(A) LENG	HARACTERIST TH: 906 ami : amino aci LOGY: linea	no acids			
(ii)		YPE: protei	† · -		* * * * * * * * * * * * * * * * * * * *	

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Met Val Arg Leu Leu Ile Phe Phe Pro Met Ile Phe Leu Glu Met

10 15

Ser Ile Leu Pro Arg Met Pro Asp Arg Lys Val Leu Leu Ala Gly Ala
20 25 30

Ser Ser Gln Arg Ser Val Ala Arg Met Asp Gly Asp Val Ile Ile Gly
40 45

Ala Leu Phe Ser Val His His Gln Pro Pro Ala Glu Lys Val Pro Glu
50 60

Arg Lys Cys Gly Glu Ile Arg Glu Gln Tyr Gly Ile Gln Arg Val Glu
65 70 80

Ala Met Phe His Thr Leu Asp Lys Ile Asn Ala Asp Pro Val Leu Leu
85 90

Pro Asn Ile Thr Leu Gly Ser Glu Ile Arg Asp Ser Cys Trp His Ser
100 100

Ser Val Ala Leu Glu Gln Ser Ile Glu Phe Ile Arg Asp Ser Leu Ile
115

Ser Ile Arg Asp Glu Lys Asp Gly Leu Asn Arg Cys Leu Pro Asp Gly
135

Gln Thr Leu Pro Pro Gly Arg Thr Lys Lys Pr Il Ala Gly Val Ile
145

GURSTITUTE SHEET

								•		•						÷ •
Gl	y-:	Pro	Gly	Ser	Ser 165	Ser	Val	Ala	Ile	Gln 170	Val	Gln	Asn	Leu	Leu 175	Gln
•				180		Gln			183		•		:	1,0		
Se	r	Asp	Lys 195	Thr	Leu	Tyr	Lys	Tyr 200	Phe	Leu	Arg	Val	Val 205	Pro	Ser	Asp
		210					215					220			•	Trp
22	5					Val 230					235	•		`		240
		/			245	Glu	. •	•		250						•
				260		Tyr		9	265					2,0	*	
			275			Arg	• •	280					205			
	-	290		,	• .	•	295	*				300				Arg
3	05	•				Gly 310		•			313					32,5
•			. ,		325		,			330	ľ				•••	
				340)	Leu	٠.		345					330		
	_		355	5			. •	360		•			303			Phe
		370)		.	•	. 375	•		,		360		•		Leu
3	85			•		390					375		•	•		400
					40!	5				41	•					
		_	•	420	0				425	•		•	-			ı Cys
			43	5		•		, 441	,						*	/ Arg
I	ys	Le: 45		u As	p Ph	e Lev	1 Ild 459	a Lys	s Sei	: Se	r Pho	460	l Gly	y Val	l Sei	r Gly

G!	lu	Glu	Val	Trp	Phe	Asp .	Glu	Lys	Gly	Asp	Ala 475	Pro	Gly	Arg	Tyr	Asp 480
46	55					4/0					-	•			Val 495	
		•			485					.,,						· .
			•	500					700	. "			•		Lys	
			515			•		720		•	•				Pro	• *
		530					333								Cys	
- 5	45					550			,						Glu	•
		V 1			565										• • •	Thr
				580					303						Ile	•
			595					600							Thr	
		610	١,				GIO				•				Val	
6	525					630			• .						8 .	Leu 640
					643											
				660	,						•					Tyr
			675	יַ נ	y -			000	, ,		•				•	Gly
•		690)		•		. 093	,								Trp
•	705	,				110	,					,				720
		•			72:	5	-				-			•	*	Tyr
				740) ·							•				ı Gly
•	Va]	L Va	I Ala 75	a Pr	Va.	l Gl	y Tyi	760	n Gly	Lei	u Let	1 Ile	76	t Se 5	r Cy	s Thr

Tyr	Tyr	Ala	Phe	Lys	Thr	Arg 775	Asn	Val	Pro	Ala	Asn 780	Phe	Asn	Glu	Ala		· ·
•	770										•					• .	•
Lys 785	Tyr	Ile	Ala	Phe	Thr 790	Met	Tyr	Thr	Thr	Cys 795	Ile	Ile	Trp	Leu	Ala 800		81
Phe	Val	Pro	Ile	Tyr 805	Phe	Gly	Ser	Asn	Tyr 810	Lys	Ile	Ile	Thr	Thr 815	Cys	*	*
Phe	Ala	Val	Ser 820	Leu	Ser	Val	Thr	Val 825	Ala	Leu	Gly	Cys	Met 830	Phe	Thr	*.	
Pro	Ĺys	Met 835	Tyr	Ile	Ile	Ile	A1a 840	Lys	Pro	Glu	Arg	Asn 845	Val	Arg	Ser		
Ala	Phe 850	Thr	Thr	Ser	Asp	Val 855	Val	Arg	Met	His	Val 860	Gly	Asp	Gly	Lys		· .
	Pro	•	Arg	Ser	Asn 870	Thr	· .	Leu	Asn	Ile 875	Phe	Arg	Arg	Lys	Lys 880	•	
		,	*	•			•		• , ,					Base	60*		
Pro	Gly	Ala	Gly	Asn 885	Ala	Lys	Lys	Arg	890	Pro	Glu	Pne	Ser	895	Ser		
Ser	Ģlņ	Cys	Pro 900	Ser	Ala	His	Ala	Glr 905	Lev i			*	- 1	ور دور ب _ر م _ه دی			
(2)	TNE	CODMA	TION	FOR	SEC	ID	NO:1	8:					•	8	•		· ·
(2)		•							,				· .· ·		,		
	(1) SE	QUEN A) I	CE C	HARA	CTEF	NISTI Dase	CS:	rs			**			÷ .		• e
		(A) I B) I	YPE:	nuc	cleid	aci	id		1 -		- '		*			
		. ((C) S	TRAN	IDEDN	vess :	sir	ngle		•							
	(15	L) MC	LECU	LE 1	YPE	: cDì	IA.					0-	:	-1		٠.	
	\	+		•		ė			. •					::	` .		
	/404 4	6 \T%	MED]	ATE	SOUI	RCE:			•	•							
	(*1.		(B) (LON	2: SI	N30											
*			2 mv11	».			0.			· .			*	٠.,			
	(1)		IUTAS (A)	NAME,	KEY	: CD	5			•		1			*	,	
	•	• • •	(B) 1	LOCA!	rion	: 46	33	198					- (· · · · · ·	• •		
٠					11	. •			7.3								
•	, (x:	i) Si	EQUE	NCE 1	DESC	RIPT	ION:	SEQ	ID	NO: 1	8:	•		. :		· .	
cc	دددد	CTCC	CGG	CAGT	GCG .	AGCA	GCTA	AG G	GCTG	GCCG	C CG	CCTC	CCTG	AGC	TCCCC	CG	60
GA	GCAG	CCGA	CCC	CTGG'	TCG	CGGC	GTTC	AC C	TCGC	CGAT	G CG	CGGT	TGGT	AGG	AGTGA	CC	120
															CCCTI		180
															CTCC		240
															TCTAC		300

		, .								14						¥0:		
		ል ሙጥ ል	AA G	AGGA(GGAG	G GG/	AGATO	TGT	GGA	ATGG	GCC /	ACCC	CGTT	G C	CTGC	rgcat	:	360
	AAciu ma cot	GYYC	CT G	CGCT	GTCC	A CAG	CGTGC	CCA	GAT	CATG	GGA	CCCA	GGC	Ț G	CTAG	GGCTA	•	420
	TACL	GAAC							000		CCG	33 & '	TG TO	C G	GG A	AG		474
	GGAG	CGGG	GC C	CAGT	ATTC	A TG	GGTCT	CTA	GGC	CTTT	-	M	et So	er G	ly L	ys _:	•	
	•		. •								. •							522
	GGA Gly	GGC Gly	TGG Trp	GCC Ala	TGG Trp	TGG ' Trp '	TGG (CC 11a	CGG Arg	CTG Leu	CCC Pro 15	CTC Leu	TGC (CTA (Leu	CTC Leu	Leu 20		:
	· •				•		•	<u>}</u>			mm <i>c</i>	dea .	A A G	CCC	AAG	GGT	·	570
	AGC	CTT	TAT	GCC	CCC	TGG	GTG (CCT	TCA	Ser	Leu	Gly	Lys	Pro	Lys	Gly		
	Ser	Leu	Tyr	Ala	Pro 25	Trp	val :	-10		30					35	· ·	•	
						let					ccc	GAC	ልጥር	ACA	CTG	GGA	4	618
<i>:</i>	CAC	CCC	CAC	ATG	AAC	TCT	ATC	CGA	ATT.	ASD	GlV	Asp	Ile	Thr	Leu	Gly		• •
	Нis	Tro.	His	Met.	Asn	Ser _.	IIG .	ary	45	,,LP		-		50	. •	-		
									:		C3.C	CCT	AAG	GCC	TGC	GGG		666
	GGC	CTG	TTT	CCC	GTC	CAC.	GGC	CGT	GGC	TCT	GAG	Gly	Lys	Ala	Cys	Gly	•	
	Gly	Leu	Phe 55	Pro	Val	HIS	GIY	60	 2			_	65	`,				
				*		٠				000	CTIC:	GAG	GCC	ATG	CTG	TTT		714
	GAĞ	CŢG	AAG	AAG	GAG	AAA	GGC Gly	ATC Tle	His	Ara	Leu	Glu	Ala	Met	Leu	Phe	•	
	Glu	Leu 70	Lys	Lys	GIU	гля	75	110				80			•			
		• -						خ د م	000	CAC	CTP A	CTG	CCC	AAC	ATC	ACG .		762
٠.	GCC	CTG	GAC	CGC	ATC	AAC	AAT Asn	Asp	Pro	ASD	Leu	Leu	Pro	Asn	Ile	Thr		
	Ala 85	Leu	Asp	Arg	TTG	90	nsu	P		*	95				٠	100	•	
			•						 	MCG	»GG	GAC	ACC	CAC	GCC	CTG	•	810
	TTG	GGC	GCC	CGC	ATT	CTG	GAC Asp	Thr	CVS	Ser	Arg	Asp	Thr	His	Ala	Leu		i
,	Leu	Gly	YIS	Arg	105	Dea	p			110		,			115		•	
		-					GTG Val	-	cès	. صند	АТС	GAG	AAG	GAC	GGC	ACG		858
	GAG	CAG	TCA	CTG	ACC	Dhe.	GTG Val	Ara	Ala	Leu	Ile	Glu	Lys	Asp	Gly	Thr	**	
				120	,		•								-	:		-
							000	ccc	cce	CCC	ATC	ATC	ACC	AAG	CCC	GAA Glu		906
	GAG	GTC	: CGC	TGG	L GGC	AGG	Arg	Gly	Pro	Pro	Ile	Ile	Thr	Lys	Pro	Glu		
(Glu	i A91	. Arg	, cya ;	GIJ	•••		140) · · · · · · ·				145					•
•							cci	.cot	• ጥርር	: GGG	AGC	TCC	GTC	TCG	ATC	ATG Met		954
	CGA	GTC	GTO	GGI	. GTC	Ile	Gly	Ala	Ser	Gly	Ser	Ser	Val	Ser	Ile	Met	,	
		150)				100				•						. •	
				<u>,</u>			- CMC	ded.	AAC	ATC	CCI	. CAG	ATC	AGC	TAT	GCC Ala	•	1002
	GTC	GCC	AA	C ATC	CIC	; CGC	Leu	Phe	Lys	Ile	Pro	Gln	Ile	Ser	ТУ	180		
•	165	ξ		•		7.0	,						• •					
						, mm/	: <u>እ</u> ርጥ	GAC	2 AAC	B AGC	c cgc	TAT	GAC	TTC	TTC	TCC Ser		1050
	TC	CAC	GC	C CCI	GAC Lat	Lau	, AGI 1 Ser	Asi	ASI	sei	r Arg	Ty:	: Asp	Phe	Phe	ser		· · · · · · · · · · · · · · · · · · ·
•	•		•		185	•			7							•		
		•	* 			. GN	~ <u>a</u> ca	ጥል(C CA	G GC(C CA	G GCC	ATC	GT	GA:	TTA		1098
	CG	g GIY	G GIN		TUI	, GM	- 200		=-			_		•				

- TOCTITUTE SHEET

	•		:			į.											
	•			200		1.			205	Ala							
	Val	Arg	Ala 215	Leu	Lys	Trp	ASN	220	Val	TCC Ser	1111		225				1146
	AGC Ser	TAC Tyr 230	GGT Gly	GAG Glu	AGT Ser	GGT Gly	GTG Val 235	GAG Glu	GCC Ala	TTT	110	CAG Gln 240	AAG Lys	TCC Ser	CGA Arg	GAG Glu	1194
,	Asn 245	Gly	Gly	Val	Cys	250	ATS	GIN	Ser	GTG Val	255	110				260	1242
	AAG Lys	ACG Thr	GGG Gly	GAG Glu	TTC Phe 265	GAC Asp	AAG Lys	ATC Ile	ATC Ile	AAA Lys 270	CGC Arg	CTA Leu	CTG Leu	GAA Glu	ACA Thr 275	TCC	1290
		GCC Ala	AGG Arg	GGT Gly 280	ATC Ile	ATC	ATC Ile	TTT Phe	GCC Ala 285	AAC Asn	GAG Glu	GAT Asp	GAC Asp	ATC Ile 290	AGG Arg	AGG Arg	1338
	GTG Val	TTG Leu	GAG Glu 295	Ala	GCT Ala	CGC	AGG Arg	GCC Ala 300	AAC	CAG Gln	ACC Thr	GGC Gly	CAC His 305		TTT	TGG Trp	1386
	ATG Met	GGT Gly 310	Ser	GAT Asp	AGC Ser	TGG Trp	GGC Gly 315	TCC	AAG Lys	AGT Ser	GCC Ala	CCT Pro 320		CTG Leu	CGC	CTT Leu	1434
	GAG Glu 325	Glu	GTG Val	GCC Ala	GAG Glu	GGC Gly 330	VIG	GTC Val	ACC	ATT Ile	CTC Leu 335	CCC	AAG Lys	AGG Arg	ATG Met	TCT Ser 340	1482
	1	•	GGG	TTC Phe	GAC Asp 345	Arg	TAC	TTC Phe	TCC	AGC Ser 350	474 J.	ACG Thr	CTG Leu	GAC Asp	AAC Asn 359	AAC Asn	1530
:	AGG Arg	CGC	AAC ASI	ATC	Trp	TTT Phe	GCC Ala	GAG Glu	Phe 365	= rrb	GAG Glu	GAC Asp	AAC Asn	TTC Phe 370	CAT His	TGC Cys	1578
	AAG Lys	TTC	AGC 1 Sei 375	- Arg	CAC His	GCG Ala	CTC Lev	AAG Lys	, Lly	G GGA E Gly	AGC Ser	CAC His	385	•	AAC Lys	TGC Cys	1626
	ACC	AAC Asi	n Arq	A GAG	G CGC	ATC	GGG Gly	GTI	GA AS	C TCG p Ser	GCC Ala	TAT		G CAG	G GAG	GGG Gly	1674
	AAC Lys	GTG		G TTO	GTG Val	ATT	s wat	GCT Ala	GT Va	G TAC	C GCC Ala		G GGG	C CAC y Hi	c GC s Al	G CTG Leu 420	1722
			C AT(G CAC	C CG1	GA(C CTC	ı Cy	9 21	C GGG	,	_	A GG. 1 Gl	À CT y Le	C TG u Cy	C CCT s Pro	1770
								~~1	T 1 1	TE.	3NC	ا عا					1274 of

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5:
	**C12P 21/06, C12N 5/00, 15/00
	C07H 15/12, 17/00, C07K 3/00
	C07K 13/00, 15/00, 17/00
	A61K 35/14

(11) International Publication Number:

WO 92/10583

(43) International Publication Date:

25 June 1992 (25.06.92)

(21) International Application Number:

PCT/US91/09422

A1

(22) International Filing Date:

12 December 1991 (12.12.91)

(30) Priority data:

626,806 12 December 1990 (12.12.90) 648,481 30 January 1991 (30.01.91) 672,007 18 March 1991 (18.03.91) US

(60) Parent Application or Grant

Filed on

(63) Related by Continuation

626,806 (CIP) 12 December 1990 (12.12.90)

(71) Applicants (for all designated States except US): ZYMOG-ENETICS, INC. [US/US]; 4225 Roosevelt Way N.E., Seattle, WA 98105 (US). THE BOARD OF REGENTS OF THE UNIVERSITY OF WASHINGTON [US/US]; Seattle, WA 98195 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): MULVIHILL, Eileen, Ranae [US/US]; 4016 Francis Avenue North, Seattle, WA 98103 (US). HAGEN, Frederick, Stamner [US/US]; 1315 Lexington Way East, Seattle, WA 98112 (US). HOUAMED, Khaled, M. [TN/US]; 730 Harvard Avenue, Seattle, WA 98102 (US). ALMERS, Wolfhard [DE/US]; 825 35th Avenue, Seattle, WA 98122 (US).

(74) Agents: PARMELEE, Steven, W. et al.; Townsend and Townsend, One Market Plaza, 2000 Steuart Tower, San Francisco, CA 94105 (US).

(81) Designated States: AT, AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI p CH, CH (European patent), CI (OAPI patent), CM (OAPI patent), CM (OAPI patent), CS, DE, DE (European patent), DK, DK (European patent), ES, ES (European patent), FI, FR (European patent), GA (OAPI patent), GB, GB (European patent), GN (OAPI patent), GR (European patent), GN (OAPI patent) ropean patent), GA (OAPI patent), GB, GB (European patent), GN (OAPI patent), GR (European patent), HU, IT (European patent), JP, KP, KR, LK, LU, LU (European patent), MC (European patent), MG, ML (OAPI patent), MN, MR (OAPI patent), MW, NL, NL (European patent), NO, PL, RO, SD, SE, SE (European patent), SN (OAPI patent), SU+,TD (OAPI patent), TG (OAPI patent), US.

Published

With international search report.

(54) Title: G PROTEIN-COUPLED GLUTAMATE RECEPTORS

(57) Abstract

the bound of the state.

Mammalian G protein-coupled glutamate receptors are identified, isolated and purified. The receptors have been cloned, sequenced and expressed by recombinant means. The receptors and antibodies thereto can be used to identify agonists and antagonists of G protein-coupled glutamate receptor mediated neuronal excitation and in methods of diagnosis.

+ DESIGNATIONS OF "SU"

Any designation of "SU" has effect in the Russian Federation. It is not yet known whether any such designation has effect in other States of the former Soviet Union.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

						,
	AT	Austria	ES	Spain	MG	Madagascar
	AU	Australia	FI	Finland	ML	Mali
	BB	Barbados	FR	France	MN	Mongolia
		Belgium	GA	Gabon	MR	Mauritania
	28	Burkina Faso	GB	United Kingdom	МŴ	Malawi
	BP		GN	Guines	NL	Netherlands
	BG ·	Bulgaria	GR	Greece	· NO	Norway
•	Đ	Benin	HU	Hungary	PĹ	Poland
	BR	Brazil	'n	Italy	RO	Romania ·
	CA	Canada			SD	Sudan
	CF	Contral African Republic	JP	Japan	SE	Sweden
	CC	Congo	KP.	Democratic People's Republic	SN	Senceal
••	CH	Switzerland	•	of Korea	su+	Soviet Union
	a	Côte d'Ivoire	· KR	Republic of Kores	TD	Chad
	СМ	- Cameroon-	LI	Llechtenstein		
	CS	Czochoslovakia	LK	Sri Lanka	TG.	Togo United States of America
	DR	Germany	· LU	Luxembourg	US	United States of Allithies
	DK	Denmark	MC	Monaco		•
	~~		* *			

G PROTEIN-COUPLED GLUTAMATE RECEPTORS

.

15

20

10

Background of the Invention

The majority of nerve cell connections are chemical synapses. A neurotransmitter is released from the presynaptic terminal, typically in response to the arrival of an action potential in the neuron, and diffuses across the synaptic space to bind to membrane receptor proteins of the postsynaptic terminal. The binding of neurotransmitters to membrane receptors is coupled either to the generation of a permeability change in the postsynaptic cell or to metabolic changes.

Neurotransmitters produce different effects according to the type of receptor to which they bind. In general, those which produce effects that are rapid in onset and brief in duration bind to receptors that act as ligand-gated ion channels, where binding almost instantly causes an ion flow across the membrane of the postsynaptic cell. Those neurotransmitters which act more like local chemical mediators bind to receptors that are coupled to intracellular enzymes, thereby producing effects that are slower in onset and more prolonged. These neurotransmitters alter the concentration of intracellular second messengers in the postsynaptic cell.

Four second messenger systems have been linked to neurotransmitter or hormone receptors and have been studied for their roles in the control of neuronal excitability. They are the adenylate cyclase/cyclic AMP-dependent pr t in kinase system, guanylat cyclase and cGMP-dependent protein kinase, th inosit l trisph sphat /diacyl glycer l-pr tein kinase C system,

25

30

15

20...

25

35

and systems which are activated by calcium i ns, such as th calcium/calm dulin-dependent prot in kinase system. Thus, binding of a transmitter to a receptor may activate, for example, adenylate cyclase, thereby increasing the intracellular concentration of cAMP. The cAMP activates protein kinases that phosphorylate proteins in the cells, which form ion channels, thereby altering the cells' electrical behavior. As with the ligand-gated ion channel transmitters, the effects can be either excitatory or inhibitory, and may affect the cell at many levels, including the pattern of gene expression. It is also believed that these chemical synapses, associated with second-messenger systems, may be involved in long-term changes that comprise the cellular basis of learning and memory.

The ligand-activated membrane receptors do not activate the second messenger systems directly, however, but via a membrane-bound protein, the GTP-binding protein (G protein), which binds GTP on the cytoplasmic surface of the cell membrane and thereby acts to couple adenylate cyclase to the membrane receptor. Neurotransmitter binding to the membrane receptor is believed to alter the conformation of the receptor protein to enable it to activate the G protein in the lipid bilayer, which then binds GTP at the cytoplasmic surface and produces a further change in the G protein to allow it to activate, e.g., an adenylate cyclase molecule to synthesize cAMP. When a ligand binds a receptor, an enzymatic cascade results as each receptor activates several molecules of G protein, which in turn activate more molecules of adenylate cyclase which convert an even larger number of ATPs to cAMP molecules, producing a substantial amplification from the initial event.

Glutamate, aspartate and their endogenous derivatives are believed to b th pred minant excitatory n urotransmitters in the v rtebrate central nervous system. (Krinjrvic, Phys. Rev. 54:418-540, 1974). R c ntly, glutamate has be n described as playing a

15

20

25

30

35

major, wid spread role in the control of neuroendocrine neurons, possibly controlling not nly the neuroendocrine system but other hypothalamic regions as well. Four major subclasses of glutamate receptors have been described but their characterization has until recently been limited to pharmacological and electrophysiological functional analyses. See generally, Hollman et al., Nature 342:643-648 (1989) and Sommer et al., Science 249:1580-1585 (1990). Three of the receptors, the quisqualate (QA/AMPA), kainate (KA), and N-methyl-Daspartate (NMDA) receptors, are believed to be directly coupled to cation-specific ion channels and thus are classified as ligand-gated ionotropic receptors. The fourth glutamate receptor binds some of the agonists of the ionotropic receptors (quisqualate and glutamate, but not AMPA) but has no shared antagonists, and is coupled to G protein. Thus, this receptor, referred to as the G protein-coupled glutamate receptor, or Gluck, is pharmacologically and functionally distinct from the other major glutamate receptors. This receptor has also been termed the metabotropic receptor.

Agonist binding to Gluck has been shown to result in the activation of a number of second messenger systems, depending on the system studied. One of the best characterized is the quisqualate activation of phospholipase C through a G protein-coupled interaction that leads to the stimulation of inositol phospholipid metabolism. This activity has been studied in systems that measure the accumulation of radiolabeled inositol monophosphate in response to stimulation by glutamate. The systems typically use brain slices from regions such as the hippocampus, striatum, cerebral cortex and hypothalamus (Nicoletti, et al., Proc. Natl. Acad. Sci. <u>USA</u> 83:1931-1935 (1986), and Nicoletti, et al., <u>J.</u> Neurochem. 46:40-46 (1986)), neuronal cultures derived from embryonic mouse and rat cerebellum, c rpus striatum and cerebral cortex (Nicoletti et al., J. Neurosci. 6:1905-1911 (1986), Sladeczek et al., Nature 317:717-719

10

15

20

25

30

35

(1985), Dumuis, et al., Nature 347:182-184 (1990), and Drejer et al., J. Neurosci. 7:2910-2916 (1987)) and rat brain synaptosomes (Recasens et al., Eur. J. Pharm. 141: 87-93 (1987), and Recasens et al., Neurochem. Int. 13:463-467 (1988)). A major disadvantage of each of these model systems is the difficulty in analyzing the pharmacological and functional activities of GlugR in an environment where other glutamate receptors and G protein-coupled receptors such as muscarinic and serotonin receptors are also present.

The Xenopus oocyte system has been used to identify GlugR as a member of the family of G proteincoupled receptors. An endogenous inositol triphosphate second messenger-mediated pathway in the oocyte allows the detection of GlugR after injection of total rat brain mRNA, in that the cocyte responds to ligand via the oocyte G protein-coupled PLC-mediated activation of a chloride channel that can be detected as a delayed, oscillatory current by voltage-clamp recording (Houamed et al., Nature 310:318-321 (1984), Gunderson et al., Proc. Royal Soc. B221:127-143 (1984), Dascal et al., Mol. Brain Res. 1:301-309 (1986), Verdoorn et al., Science 238:1114-1116 (1987), Sugiyama et al., Nature 325:531-533 (1987), Hirono et al., Neuros. Res. 6:106-114 (1988), Verdoorn and Dingledine, Mol. Pharmacol. 34:298-307 (1988), and Sugiyama et al., Neuron 3:129-132 (1989)). Injection of region-specific brain mRNA and of size fractionated mRNA have suggested that GlugR may be a large mRNA (6-7 kb) and that it is enriched in the cerebellum (Fong et al., Synapse 2:657-665 (1988) and Horikoshi et al., Neurosci. Lett. 105:340-343 (1989)).

There remains considerable need in the art for isolated and purified Glu_cR, as well as systems capable of expressing Glu_cR separate from other neurotransmitter receptors. Further, it would b desirabl t specifically identify the presence of Glu_cR in cells and tissues, thereby avoiding the time-consuming, complex and nonspecific functional electrophysiological and

15

20

25

30

35

pharmac l gical assays. It will also be d sirable t screen and dev lop new agonists and/or antag nists specific for Glu_cR , but to date this has not been practical. Quite surprisingly, the present invention fulfills these and other related needs.

Summary of the Invention

The present invention provides isolated and substantially pure preparations of mammalian G protein-coupled glutamate receptors and fragments thereof. In preferred embodiments the receptors are coupled to a G protein in vertebrate cells, bind glutamate and quisqualate and thereby activate phospholipase C, and are capable of stimulating inositol phospholipid metabolism. Having provided such receptors in isolated and purified form, the invention also provides antibodies to the receptors, in the form of antisera and/or monoclonal antibodies.

In another aspect the invention provides the ability to produce the mammalian G protein-coupled glutamate receptors and polypeptides or fragments thereof by recombinant means, preferably in cultured eukaryotic cells. The expressed receptors or fragments may or may not have the biological activity of corresponding native receptors, and may or may not be coupled to a G protein in the cell used for expression. Accordingly, isolated and purified polynucleotides are described which code for the receptors and fragments thereof, where the polynucleotides may be in the form of DNA, such as cDNA, or RNA. Based on these sequences probes may be used to hybridize and identify these and related genes which encode mammalian G protein-coupled glutamate receptors. The probes may be full length cDNA or as small as from 14 to 25 nucleotide, more often though from about 40 to ab ut 50 r m re nucleotid s.

In relat d emb dim nts the inventi n c ncerns

DNA c nstructs which c mprise a transcriptional pr m t r,

a DNA sequ nce which enc des the recept r r fragment,

15

20

25

30

35

and a transcriptional t rminator, each operably linked for expression of the rec ptor. For expression the construct may also contain at least one signal sequence. The constructs are preferably used to transform or transfect eukaryotic cells, more preferably mammalian cells which do not express endogenous G protein-coupled glutamate receptors. When bound by an appropriate ligand such as glutamate or quisqualate, the receptor may activate phospholipase C in the host cell via coupling to G protein. Further, for large scale production the expressed receptor may also be isolated from the cells by, for example, immunoaffinity purification.

Cells which express the G protein-coupled glutamate receptors may also be used to identify compounds which can alter the receptor-mediated metabolism of a eukaryotic cell. Compounds may be screened for binding to the receptor, and/or for effecting a change in receptor-mediated metabolism in the host cell. Agonists and/or antagonists of the G protein-coupled glutamate receptors may also be screened in cell-free systems using purified receptors or binding fragments thereof for the effect on ligand-receptor interaction, or using reconstituted systems such as micelles which also provide the ability to assess metabolic changes.

In yet other embodiments the invention relates to methods for diagnosis, where the presence of a mammalian G protein-coupled glutamate receptor in a biological sample may be determined. For example, a monospecific antibody which specifically binds a G protein-coupled glutamate receptor is incubated with the sample under conditions conducive to immune complex formation, which complexes are then detected, typically by means of a label such as an enzyme, fluorophore, radionuclide, chemiluminescer, particle, or a second labeled antibody. Thus, means are provided for immunohistochemical staining of tissues, including brain tissues, for the subject receptors.

15

20

25

30

35

7

Brief Description of the Figures

Figure 1 illustrates the construction of plasmid pVEGT, where Fig. 1A shows the construction of pVEG, Fig. 1B shows the construction of pVEG' and Fig. 1C shows pVEGT'. Symbols used are T7 pro, the T7 promoter; T1 and T2, synthetic and native T7 terminators, respectively; M13, M13 intergenic region; the parentheses indicate a restriction site destroyed in vector construction; and pA is the Aspergillus niger polyadenylate sequence.

Figure 2 illustrates representative responses from voltage-clamp assays of oocytes injected with RNA from positive pools.

Figure 3 illustrates a partial restriction map of clone 45-A.

Figure 4 illustrates the cloning of the receptor cDNA present in clone 45-A into Zem228R.

Figure 5 illustrates the DNA sequence and deduced amino acid sequence of clone 45-A (corresponding to Sequence ID Nos. 1 and 2). Numbers below the line refer to amino acid sequence, numbers above the line refer to nucleotide number. Putative transmembrane domains have been overlined, and putative N-linked glycosylation sites are indicated by closed circles.

Figure 6 illustrates a representative dose response curve for varying concentrations of L-glutamic acid. Error bars, where larger than the symbols, represent SEM.

Figure 7 illustrates the DNA sequence and deduced amino acid sequence of a subtype 1b glutamate receptor clone (Sequence ID Nos. 16 and 17). Numbers below the line refer to amino acid sequence. Numbers above the line refer t nucleotid sequence.

Figure 8 illustrates the DNA sequ nc and deduced amin acid sequenc of a subtype 2a glutamate

SUBSTITUTE SHEET

10

15

20

25

30

35

receptor clone (Sequence ID Nos. 18 and 19). Numb rs below th line r fer to amino acid sequence. Numbers above the line refer to nucleotide sequence.

Figure 9 illustrates the DNA sequence of a partial subtype 2b glutamate receptor clone (Sequence ID No. 20). Numbers refer to the nucleotide sequence.

Description of the Specific Embodiments

Glu_GR is a family of G protein-coupled membrane receptors for the neurotransmitter glutamate. As glutamate has been described as having a major role in the control of neurons, particularly neuroendocrine neurons, Glu_GR may play a critical role in effectuating such control. Consequently, the development of agonists and antagonists of the Glu_GR-ligand interaction and Glu_GR-mediated metabolism is of great interest.

The present invention presents the means to identify agonists and antagonists of the GlugR-ligand interaction by providing isolated GluGR. The term "GlugR" refers to any protein either derived from a naturally occurring Glu_GR , or which shares significant structural and functional characteristics peculiar to a naturally occurring GluGR. Such a receptor may result when regions of a naturally occurring receptor are deleted or replaced in such a manner as to yield a protein having a similar function. Homologous sequences, allelic variations, and natural mutants; induced point, deletion, and insertion mutants; alternatively expressed variants; proteins encoded by DNA which hybridize under high or low stringency conditions to nucleic acids which encode naturally occurring GluGR-encoding nucleic acids; proteins r trieved from naturally occurring materials; and clos ly related prot ins r trieved by antisera directed against GluGR proteins are also included.

WO 92/10583 PCT/US91/09422

9

analog, or chimeric Glu_cR as generally described in U.S. Pat. No. 4,859,609, incorporated by ref r nce herein. The molecule may be chemically synthesized or may occur in nature. Ligands may be grouped into agonists and antagonists. Agonists are those molecules whose binding to a receptor induces the response pathway within a cell. Antagonists are those molecules whose binding to a receptor blocks the response pathway within a cell.

10

15

20

25

30

35

By "isolated" Glu_cR is meant to refer to a Glu_cR which is in other than its native environment such as a neuron, including, for example, substantially pure Glu_cR as defined hereinbelow. More generally, isolated is meant to include a Glu_cR as a heterologous component of a cell or other system. For example, a Glu_cR may be expressed by a cell transfected with a DNA construct which encodes the Glu_cR, separated from the cell and added to micelles which contain other selected receptors. In another example described below, a Glu_cR is expressed by a cell which has been co-transfected with a gene encoding muscarinic receptor. Thus, in this context, the environment of the isolated Glu_cR is not as it occurs in its native state, particularly when it is present in a system as an exogenous component.

The invention provides cloned Glu_cR coding sequences which are capable of expressing Glu_cR proteins. Complementary DNA encoding Glu_cR may be obtained by constructing a cDNA library from mRNA from, for example, brain tissue. The library may be screened by transcribing the library and injecting the resulting mRNA into occytes and detecting, by functional assays, those occytes which express the Glu_cR. Alternatively, the clones may be screened with a complementary labeled oligonucleotide probe.

The present invention relates to successfully isolating a cDNA encoding a Glu_cR. Functional cloning of Glu_cR was accomplished by substantial modifications and improvem nts to a number of cDNA cl ning and molecular biology techniqus. Initially, an enriched source f

15

20

25

30

35

Glu R mRNA pr pared by sucros gradi nt centrifugation f >4kb length rat cerebellum poly(A)+ mRNA was used as template for cDNA synthesis. Further, a cDNA cloning vector that was employed included a poly(A) tail, thereby increasing by 40-fold the translational efficiency of the transcription product of the cDNA insert and a polylinker site to allow the directional cloning of the cDNA into the vector between the promoter and the poly(A) tail. Vector construction for directional cloning is described in co-pending U.S.S.N. 07/320,191, incorporated herein by The cDNA cloning vector also was used with reference. two transcriptional terminators, in tandem, following the poly(A) sequences, efficiently generating a unit length transcript product without non-coding plasmid or viral sequences, and without requiring a restriction endonuclease to linearize the DNA template (a standard practice that will often prevent functional cloning strategies from working due to the presence of the endonuclease site within the coding region of the cDNA). The cDNA synthesis strategy maximized insert size and recreation of the 5' ends of the cDNA's, without introduction of homopolymer tails. cDNA inserts were size-selected to be greater than 4 kb in length before insertion into the vector. A library of 106 cDNA inserts in pools of 100,000 was replica plated to reduce the number of amplification steps in the fractionation of sequentially smaller pools. Moreover, ml muscarinic cDNA (another G protein-coupled receptor coupled to phosphoinositol metabolism) template was included in transcription reactions of the subfractionated pools so that before injection the in vitro transcripts from each pool could be assayed by Northern analysis to assess relative quantity and quality of the mRNA, and by voltage-clamp of oocytes as an internal positive control for each occyte not r sponding to quisqualate r Th inclusion f a dilution of SEAP-VEGT (a secr ted form of alkaline phosphatase) t mplate in transcriptions was also employed s that occytes selected

15

20

25

30

35

for voltage-clamp analysis were those synthesizing higher lev ls of th co-inject d Glu_CR mRNA. And further, low noise electrical recording techniques were used to monitor the small signals initially generated from rare transcripts.

The above-described methods were used to isolate a cDNA clone encoding a Glu_GR designated "subtype la." Oligonucleotide probes based on the sequence of the subtype la clone were used to probe additional brain and cerebellum cDNA libraries. These libraries yielded clones encoding additional subtypes, which were designated 1b, 2a and 2b.

With the Gluck and cDNA clones thereof provided herein, nucleotide and amino acid sequences may be determined by conventional means, such as by dideoxy sequencing. See generally, Sambrook et al., Molecular Cloning, A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989, incorporated by reference herein. Genomic or cDNA sequences encoding Gluck and homologous receptors of this family may be obtained from libraries prepared from other mammalian species according to well known procedures. For instance, using oligonucleotide probes from rodent Gluck, such as whole length cDNA or shorter probes of at least about fourteen nucleotides to twenty-five or more nucleotides in length; often as many as 40 to 50 nucleotides, DNA sequences encoding Glu_R of other mammalian species, such as lagomorph, avian, bovine, porcine, murine, etc. may be obtained. If partial clones are obtained, it is necessary to join them in proper reading frame to produce a full length clone, using such techniques as endonuclease cleavage, ligation and loopout mutagenesis.

A DNA sequence encoding Gluck is inserted into a suitable expression vector, which in turn is used to transfect ukaryotic cells. Expr ssion v ct rs for us in carrying out the pr sent inv ntion will comprise a

10

15

20

25

30

35

promoter capable of directing the transcription of a cloned DNA and a transcriptional terminator.

To direct proteins of the present invention for transport to the plasma membrane, at least one signal sequence is operably linked to the DNA sequence of interest. The signal sequence may be derived from the GlugR coding sequence, from other signal sequences described in the art, or synthesized de novo.

Host cells for use in practicing the present invention include mammalian, avian, plant, insect and fungal cells, but preferably mammalian cells. Fungal cells, including species of yeast (e.g., Saccharomyces spp., particularly S. cerevisiae, Schizosaccharomyces spp.) or filamentous fungi (e.g., Aspergillus spp., Neurospora spp.) may be used as host cells within the present invention. Suitable yeast vectors for use in the present invention include YRp7 (Struhl et al., Proc. Natl. Acad. Sci. USA 76: 1035-1039, 1978), YEp13 (Broach et al., <u>Gene</u> 8: 121-133, 1979), POT vectors (Kawasaki et al, U.S. Patent No. 4,931,373, which is incorporated by reference herein), pJDB249 and pJDB219 (Beggs, Nature 275:104-108, 1978) and derivatives thereof. Such vectors will generally include a selectable marker, which may be one of any number of genes that exhibit a dominant phenotype for which a phenotypic assay exists to enable transformants to be selected. Preferred selectable markers are those that complement host cell auxotrophy, provide antibiotic resistance or enable a cell to utilize specific carbon sources, and include LEU2 (Broach et al., ibid.), <u>URA3</u> (Botstein et al., <u>Gene</u> 8: 17, 1979), <u>HIS3</u> (Struhl et al., ibid.) or POT1 (Kawasaki et al., ibid.). Another suitable selectable marker is the CAT gene, which confers chloramphenicol resistance on yeast cells.

Additional vectors, promoters and terminators for use in xpressing the recept r of th inv ntion in y ast are w 11 km wn in the art and ar reviewed by, for example, Emr, Meth. Enzymol. 185:231-279, (1990), incorporat d herein by referenc. The receptors of the

invention may be expressed in Aspergillus spp. (McKnight and Upshall, d scribed in U.S. Patent 4,935,349, which is incorporated herein by reference). Useful promoters include those derived from Aspergillus nidulans glycolytic genes, such as the ADH3 promoter (McKnight et al., EMBO J. 4:2093-2099, 1985) and the tpiA promoter. An example of a suitable terminator is the ADH3 terminator (McKnight et al., ibid.). Techniques for transforming fungi are well known in the literature, and have been described, for instance by Beggs (ibid.), Hinnen et al. (Proc. Natl. Acad. Sci. USA 75:1929-1933, 1978), Yelton et al. (Proc. Natl. Acad. Sci. USA 81:1740-1747, 1984), and Russell (Nature 301:167-169, 1983) each of which are incorporated herein by reference.

A variety of higher eukaryotic cells may serve as host cells for expression of the Glu_R, although not all cell lines will be capable of functional coupling of the receptor to the cell's second messenger systems. Cultured mammalian cells, such as BHK, CHO, Y1 (Shapiro et al., TIPS Suppl. 43-46 (1989)), NG108-15 (Dawson et al., Neuroscience Approached Through Cell Culture, Vol. 2, pages 89-114 (1989)), N1E-115 (Liles et al., J. Biol. Chem. 261:5307-5313 (1986)), PC 12 and COS-1 (ATCC CRL 1650) are preferred. Preferred BHK cell lines are the tk ts13 BHK cell line (Waechter and Baserga, Proc. Natl. Acad. Sci. USA 79:1106-1110 (1982)) and the BHK 570 cell line (deposited with the American Type Culture) Collection, 12301 Parklawn Dr., Rockville, MD. under accession number CRL 10314). A tk BHK cell line is available from the ATCC under accession number CRL 1632.

Mammalian expression vectors for use in carrying out the present invention will include a promoter capable of directing the transcription of a cloned gene or cDNA. Preferred promoters include viral promot rs and cellular promot rs. Viral prom t rs include the immediat arly cytomegalovirus promot r (Boshart et al., Cell 41: 521-530, 1985) and the SV40 promot r (Subramani et al., Mol. Cell. Biol. 1: 854-864,

10

15

25

30

metallothionein-1 promoter (Palmiter et al., U.S. Pat nt No. 4,579,821), a mouse V_K promoter (Bergman et al., Proc. Natl. Acad. Sci. USA 81: 7041-7045, 1983; Grant et al., Nuc. Acids Res. 15: 5496, 1987) and a mouse V_H promoter (Loh et al., Cell 33: 85-93, 1983). A particularly preferred promoter is the major late promoter from Adenovirus 2 (Kaufman and Sharp, Mol. Cell. Biol. 2: 1304-13199, 1982). Such expression vectors may also contain a set of RNA splice sites located downstream from the promoter and upstream from the DNA sequence encoding the peptide or protein of interest. Preferred RNA splice sites may be obtained from adenovirus and/or immunoglobulin genes.

Also contained in the expression vectors is a polyadenylation signal located downstream of the coding sequence of interest. Polyadenylation signals include the early or late polyadenylation signals from SV40 (Kaufman and Sharp, ibid.), the polyadenylation signal from the Adenovirus 5 E1B region and the human growth hormone gene terminator (DeNoto et al., Nuc. Acids Res. 9: 3719-3730, 1981). The expression vectors may include a noncoding viral leader sequence, such as the Adenovirus 2 tripartite leader, located between the promoter and the RNA splice sites. Preferred vectors may also include enhancer sequences, such as the SV40 enhancer and the mouse μ enhancer (Gillies, Cell 33: 717-728, 1983). Expression vectors may also include sequences encoding the adenovirus VA RNAs.

Cloned DNA sequences may be introduced into cultured mammalian cells by, for example, calcium phosphate-mediated transfection (Wigler et al., Cell 14: 725, 1978; Corsaro and Pearson, Somatic Cell Genetics 7: 603, 1981; Graham and Van der Eb, Virology 52: 456, 1973.) Oth r techniques for introducing cloned DNA s quences int mammalian cells, such as electroporation (N umann et al., EMBO J. 1: 841-845, 1982), may also be used. In order to identify cells that have int grated

15

10

20

25

30

the cloned DNA, a s lectable marker is gen rally introduc d into the cells along with the gene or cDNA of interest. Preferred selectable markers for use in cultured mammalian cells include genes that confer resistance to drugs, such as neomycin, hygromycin, and methotrexate. The selectable marker may be an amplifiable selectable marker. Preferred amplifiable selectable markers are the DMFR gene and the neomycin resistance gene. Selectable markers are reviewed by Thilly (Mammalian Cell Technology, Butterworth Publishers, Stoneham, MA, which is incorporated herein by reference). The choice of selectable markers is well within the level of ordinary skill in the art.

Selectable markers may be introduced into the cell on a separate plasmid at the same time as the gene of interest, or they may be introduced on the same plasmid. If on the same plasmid, the selectable marker and the gene of interest may be under the control of different promoters or the same promoter, the latter arrangement producing a dicistronic message. Constructs of this type are known in the art (for example, Levinson and Simonsen, U.S. Patent No. 4,713,339). It may also be advantageous to add additional DNA, known as "carrier DNA" to the mixture which is introduced into the cells.

for a period of time, typically 1-2 days, to begin expressing the DNA sequence(s) of interest. Drug selection is then applied to select for growth of cells that are expressing the selectable marker in a stable fashion. Transfected cells may also be selected in the presence of antagonist to inhibit the activity of the receptor. Suitable antagonists in this context include D, L, 2-amino-3-phosphonopropionate. For cells that have been transfected with an amplifiable selectable marker the drug c ncentration may be increased in a st pwise manner t select for increased c py number of the cl ned sequences, thereby increasing expression l vels.

10

15

20

25

30

Promoters, terminators and methods suitable for introducing expression vectors incoding recombinant Gluck into plant, avian and insect cells are known in the art. The use of baculoviruses, for example, as vectors for expressing heterologous DNA sequences in insect cells has been reviewed by Atkinson et al. (Pestic. Sci. 28: 215-224,1990). The use of Agrobacterium rhizogenes as vectors for expressing genes in plant cells has been reviewed by Sinkar et al. (J. Biosci. (Banglaore) 11: 47-58, 1987).

Host cells containing DNA constructs of the present invention are then cultured to produce recombinant Gluck. The cells are cultured according to accepted methods in a culture medium containing nutrients required for growth of mammalian or other host cells. A variety of suitable media are known in the art and generally include a carbon source, a nitrogen source, essential amino acids, vitamins, minerals and growth factors. The growth medium will generally select for cells containing the DNA construct by, for example, drug selection or deficiency in an essential nutrient which is complemented by the selectable marker on the DNA construct.

be detected by several methods. By transfecting cells with an expression vector containing expression units for both the Gluck and a reporter gene (e.g. luciferase), the activity of the reporter gene provides an indicator of expression of the cotransfected Gluck clone. By including one or more cyclic AMP response elements (CRE) in the reporter gene expression unit, clones encoding receptors coupled to either the stimulation or inhibition of the second messenger adenylate cyclase can be identified by a change in reporter gene expression in response to added ligand. DNA constructs comprising a linked CRE and report r gene ar known in the art. See, for example, Mellon et al., Proc. Natl. Acad. Sci. USA 86: 4887-4891 (1989), incorporated h rein by reference. Cell lines

10

15

20

25

30 -

35

expressing functional rec ptors can also be d tect d by electrophysiological measurements of agonist-induced channel activity. Receptor activity can also be assayed by measuring cytosolic free calcium concentrations in transfected cells. See, for example, Thastrup et al., Proc. Natl. Acad. Sci. USA 87: 2466-2470 (1990) and Picard et al., Science 247: 327-329 (1990), which are incorporated herein by reference. A preferred method for measuring cytosolic free calcium is by scanning cells with a fluorescent microscope coupled to a video camera. The cells are injected with a fluorescent Ca^{2*} indicator (e.g. Fluo-3 or Fura-2, Molecular Probes, Inc., Eugene, OR) and exposed to ligand.

The Glu_GR produced according to the present invention may be purified from the recombinant expression systems or other sources using purification protocols that employ techniques generally available to those skilled in the art. The most convenient sources for obtaining large quantities of Glu_GR are cells which express the recombinant receptor. However, other sources, such as tissues, particularly brain tissues of the cerebellum which contain Glu_GR, may also be employed.

Purification may be achieved by conventional chemical purification means, such as liquid chromatography, lectin affinity chromatography, gradient centrifugation, and gel electrophoresis, among others. Methods of protein purification are known in the art (see generally, Scopes, R., Protein Purification, Springer-Verlag, NY (1982), which is incorporated herein by reference) and may be applied to the purification of the GluR and particularly the recombinantly produced Glu R described herein. In a preferred embodiment immunoaffinity chromatography is employed using antibodies directed against GlugR as herein described. In another method of purification, a recombinant gene encoding Gluck or portions thereof can be modified at the amin terminus, just b hind a signal p ptide, with a sequence coding for a small hydrophilic p ptide, such as

described in U.S. Patent Nos. 4,703,004 and 4,782,137, incorporated herein by reference. Specific antib dies for the peptide facilitate rapid purification of Glu_cR , and the short peptide can then be removed with enterokinase.

Thus, as discussed above, the present invention provides Glu_GR isolated from its natural cellular environment, substantially free of other G protein-coupled glutamate receptors. Purified Glu_GR is also provided. Substantially pure Glu_GR of at least about 50% is preferred, at least about 70-80% more preferred, and 95-99% or more homogeneity most preferred, particularly for pharmaceutical uses. Once purified, partially or to homogeneity, as desired, the recombinant Glu_GR or native Glu_GR may then be used to generate antibodies, in assay procedures, etc.

In another aspect, the invention concerns polypeptides and fragments of Glu_R. Polypeptides and fragments of Gluck may be isolated from recombinant expression systems or may be synthesized by the solid phase method of Merrifield, Fed. Proc. 21:412 (1962), Merrifield, J. Am. Chem. Soc. 85:2149 (1963), or Barany and Merrifield, in The Peptides, vol. 2, pp. 1-284 (1979) Academic Press, NY, each of which are incorporated herein by reference, or by use of an automated peptide synthesizer. By "polypeptides" is meant a sequence of at least about 3 amino acids, typically 6 or more, up to 100-200 amino acids or more, including entire proteins. For example, the portion(s) of GlugR proteins which bind ligand may be identified by a variety of methods, such as by treating purified receptor with a protease or a chemical agent to fragment it and determine which fragment is able to bind to labeled glutamate in a ligand blot. Polypeptides may then be synthesized and used as antig n, to inhibit ligand-GluR interaction, etc. should be understood that as used herein, reference to Gluck is meant to include the proteins, polypeptides, and fragments thereof unl ss the context indicat s oth rwise.

20

5

10

15

25

30

15

20

25

30

35

In another aspect, the invention provides means for regulating the Glu_CR-ligand int raction, and thus treating, therapeutically and/or prophylactically, a disorder which can be linked directly or indirectly to a Glu_CR or to its ligands, such as glutamate and other endogenous excitatory amino acids. By virtue of having the receptors of the invention, agonists or antagonists may be identified which stimulate or inhibit the interaction of ligand with a Glu_CR. With either agonists or antagonists the metabolism and reactivity of cells which express the receptor are controlled, thereby providing a means to abate or in some instances prevent the disease of interest.

Thus, the invention provides screening procedures for identifying agonists or antagonists of events mediated by the ligand-Gluck interaction. Such screening assays may employ a wide variety of formats, depending to some extent on which aspect of the ligand/receptor/G protein interaction is targeted. For example, such assays may be designed to identify compounds which bind to the receptor and thereby block or inhibit interaction of the receptor with the ligand. Other assays can be designed to identify compounds which can substitute for ligand and therefore stimulate Gluck-mediated intracellular pathways. Yet other assays can be used to identify compounds which inhibit or facilitate the association of Gluck to G protein and thereby mediate the cellular response to Gluck ligand.

In one functional screening assay, the initiation of fertilization activation events are monitored in eggs which have been injected with, e.g., mRNA which codes for GlugR and subsequently exposed to selected compounds which are being screened, in conjunction with or apart from an appropriate ligand. See generally, Kline et al., Science 241:464-467 (1988), incorporated herein by reference. Occytes injected with mRNA coding for GlugR can also be assayed by measurem nt of fr cytos lic Ca^{2*} as described abov .

10

15

20

25

30

35

Another screening assay is based on th use of mammalian cell lines which express Glu R functionally coupled to a mammalian G protein. In this assay, compounds are screened for their relative affinity as receptor agonists or antagonists by comparing the relative receptor occupancy to the extent of ligand induced stimulation or inhibition of second messenger metabolism. For example, activation of phospholipase C leads to increased inositol monophosphate metabolism. Means for measuring inositol monophosphate metabolism are generally described in Subers and Nathanson, J. Mol. Cell, Cardiol. 20:131-140 (1988), incorporated herein by reference. As noted previously, receptor subtypes that are coupled to the stimulation or inhibition of the second messenger adenylate cyclase can be used in assay systems wherein reporter gene (e.g. luciferase) activity is linked to receptor-ligand interactions.

The screening procedure can be used to identify reagents such as antibodies which specifically bind to the receptors and substantially affect their interaction with ligand, for example. The antibodies may be monoclonal or polyclonal, in the form of antiserum or monospecific antibodies, such as purified antiserum or monoclonal antibodies or mixtures thereof. For administration to humans, e.g., as a component of a composition for in vivo diagnosis or imaging, the antibodies are preferably substantially human to minimize immunogenicity and are in substantially pure form. By substantially human is meant generally containing at least about 70% human antibody sequence, preferably at least about 80% human, and most preferably at least about 90-95% or more of a human antibody sequence to minimize immunogenicity in humans.

Antibodies which bind GlugR may be produced by a vari ty of means. The production of non-human antisera or monoclonal antibodies, e.g., murine, lagomorpha, equine, tc. is well known and may be accomplished by, for xample, immunizing the animal with the receptor

10

15

20

25

30

35

molecule or a preparati n c ntaining a desired portion of the receptor molecule, such as that domain or domains which contributes to ligand binding. Receptor subtypespecific antibodies can be generated by immunizing with specific peptides. Small peptides (e.g., about 14-20 amino acids) can be coupled to keyhole limpet hemocyanin, for example, to enhance immunogenicity. For the production of monoclonal antibodies, antibody producing cells obtained from immunized animals are immortalized and screened, or screened first for the production of antibody which binds to the receptor protein and then immortalized. As the generation of human monoclonal antibodies to human Gluck antigen may be difficult with conventional techniques, it may be desirable to transfer antigen binding regions of the non-human antibodies, e.g. the F(ab'), or hypervariable regions, to human constant regions (Fc) or framework regions by recombinant DNA techniques to produce substantially human molecules. Such methods are generally known in the art and are described in, for example, U.S. Patent No. 4,816,397 and EP publications 173,494 and 239,400, which are incorporated herein by reference. Alternatively, one may isolate DNA sequences which code for a human monoclonal antibody or portions thereof that specifically bind to the human receptor protein by screening a DNA library from human B cells according to the general protocol outlined by Huse et al., Science 246:1275-1281 (1989), incorporated herein by reference, and then cloning and amplifying the sequences which encode the antibody (or binding fragment) of the desired specificity.

In other embodiments, the invention provides screening assays conducted in vitro with cells which express the receptor. For example, the DNA which encodes the receptor or selected portions thereof may be transfected into an established cell line, e.g., a mammalian cell lin such as BHK or CHO, using procedures known in the art (see, e.g., Sambrook et al., Molecular Cloning, A Laboratory Manual, 2d d., Cold Spring Harbor

Laboratory Pr ss, Cold Spring Harbor, N.Y., 1989, which is incorporated herein by reference). The receptor is then expressed by the cultured cells, and selected agents are screened for the desired effect on the cell, separately or in conjunction with an appropriate ligand such as glutamate or quisqualate. Means for amplifying nucleic acid sequences which may be employed to amplify sequences encoding the receptor or portions thereof are described in U.S. Pat. Nos. 4,683,195 and 4,683,202, incorporated herein by reference.

In yet another aspect, the screening assays provided by the invention relate to transgenic mammals whose germ cells and somatic cells contain a nucleotide sequence encoding Glu_cR protein or a selected portion of the receptor which, e.g., binds ligand, GTP binding protein, or the like. There are several means by which a sequence encoding, for example, the human Glu_cR may be introduced into a non-human mammalian embryo, some of which are described in, e.g., U.S. Patent No. 4,736,866, Jaenisch, Science 240-1468-1474 (1988) and Westphal et al., Annu. Rev. Cell Biol. 5:181-196 (1989), which are incorporated herein by reference. The animal's cells then express the receptor and thus may be used as a convenient model for testing or screening selected agonists or antagonists.

In another aspect the invention concerns diagnostic methods and compositions. By means of having the Gluck molecule and antibodies thereto, a variety of diagnostic assays are provided. For example, with antibodies, including monoclonal antibodies, to Gluck, the presence and/or concentration of receptor in selected cells or tissues in an individual or culture of interest may be determined. These assays can be used in the diagnosis and/or treatment of diseases such as, for example, c rebral ischemia, Parkinsons, senile dementia and other cognitive disorders, Huntington's chorea, amyotrophic lateral scl rosis, em sis, migrain, and others.

WO 92/10583 PCT/US91/09422

5

10

15

20

25

30

35

23

Numerous types of immunoassays are available and are known to those skilled in the art, e.g., competitive assays, sandwich assays, and the like, as generally described in, e.g., U.S. Pat. Nos. 4,642,285; 4,376,110; 4,016,043; 3,879,262; 3,852,157; 3,850,752; 3,839,153; 3,791,932; and Harlow and Lane, Antibodies, A Laboratory Manual, Cold Spring Harbor Publications, N.Y. (1988), each incorporated by reference herein. assay format GlugR is identified and/or quantified by using labeled antibodies, preferably monoclonal antibodies which are reacted with brain tissues, e.g., cortex, striatum, hippocampus, cerebellum, and determining the specific binding thereto, the assay typically being performed under conditions conducive to immune complex formation. Unlabeled primary antibody can be used in combination with labels that are reactive with primary antibody to detect the receptor. For example, the primary antibody may be detected indirectly by a labeled secondary antibody made to specifically detect the primary antibody. Alternatively, the anti-GlugR antibody can be directly labeled. A wide variety of labels may be employed, such as radionuclides, particles (e.g., gold, ferritin, magnetic particles, red blood cells), fluorophores, chemiluminescers, enzymes, enzyme substrates, enzyme cofactors, enzyme inhibitors, ligands (particularly haptens), etc.

The Glu_CR DNA may be directly detected in cells with a labeled Glu_CR DNA or synthetic oligonucleotide probe in a hybridization procedure similar to the Southern or dot blot. Also, the polymerase chain reaction (Saiki et al., <u>Science</u> 239:487 (1988), and U.S. Pat. No. 4,683,195) may be used to amplify DNA sequences, which are subsequently detected by their characteristic size on agarose gels, Southern blots of these gels using Glu_CR DNA or a oligonucleotide probe, or a dot blot using similar probes. The probes may comprise from about 14-nucleotides to about 25 or more nucleotides, pref rably, 40 to 60 nucleotid s, and in some instanc s a substantial

15

20

25

30

35

portion or ev n the entire cDNA of GlugR may be used. The prob s are labeled with a detectable signal, such as an enzyme, biotin, a radionuclide, fluorophore, chemiluminescer, paramagnetic particle, etc.

Kits can also be supplied for use with the receptor of the subject invention in the detection of the presence of the receptor or antibodies thereto, as might be desired in the case of autoimmune disease. Thus, antibodies to $Glu_{c}R$, preferably monospecific antibodies such as monoclonal antibodies, or compositions of the receptor may be provided, usually in lyophilized form in a container, either segregated or in conjunction with additional reagents, such as anti-antibodies, labels, gene probes, polymerase chain reaction primers and polymerase, and the like.

The following examples are offered by way of illustration, not by limitation.

EXAMPLE I

Preparation of Gluck enriched mRNA

Total RNA was prepared from the cerebellum of rats using guanidine isothiocyanate (Chirgwin et al. Biochemistry 18:52-94 (1979)) and CsCl centrifugation. Poly(A)+ RNA was isolated using oligo d(T) cellulose chromatography. After 2 rounds of chromatography on oligo d(T) cellulose the RNA (800 µg) was divided into two aliquots and layered over 10-40% linear sucrose gradients in tubes for an SW 28 rotor. The gradients were centrifuged for 28 hours at 25,000 rpm to pellet RNA greater than 4 kb in size. The enriched RNA was injected into frog oocytes and assayed for the presence of the Glu_GR.

15

20

25

30

35

Injection of occytes and voltage-clamp assay of GlugR activity

Occytes were prepared from ovarian lobes that were surgically removed from anesthetized <u>Xenopus</u> females. The ovarian lobes were washed, pulled apart into small clumps and dissociated by treatment with collagenase for 2-3 hours at 20°C with constant, gentle agitation. The dissociation and defolicularization of the occytes is completed manually after removal of the collagenase. Occytes that were judged healthy and greater than 1 mm in diameter were transferred to a 50 mm sterile tissue culture dish and incubated in sterile, antibiotic-supplemented Barth's medium (88 mM NaCl, 1mM KCl, 0.82 mM MgSO₄, 0.33 mM Ca(NO₃)₂, 0.41 mM CaCl₂, 2.4 mM NaHCO₃, 10 mM HEPES, pH 7.4, 0.1 mg/ml gentamicin, 0.01 mg/ml penicillin, 0.01 mg/ml streptomycin, 0.5 mM theophylline, and 2.5 mM Na pyruvate) at 19°C.

Injection pipettes were pulled from hard glass tubing (Drummond) on a modified 700C Kopf vertical puller. The tip was broken and bevelled using a List Medical microforge. Tip diameters of the pipettes ranged from 20-30 mM. Injection pipettes were made RNase free by heating to 285°C overnight.

Following overnight incubation, healthy oocytes were selected for injection. RNA, which was stored at -70°C in DEPC-treated water, was thawed and centrifuged at 15,000 g for five minutes. Injection was performed using a modified pipetting device (Drummond). After injection, the oocytes were incubated in fresh, sterile Barth's medium which was changed daily, and unhealthy oocytes were removed.

Voltage-clamp assays were carried out on injected oocytes which were each placed in a small chamber of approximately 500 μ l in volume and which was continuously perfused with standard frog Ringer's (115 mM NaCl, 2.5 mM KCl, 1.8 mM CaCl₂, 10 mM HEPES, pH 7.2) at 1-6 ml/min. The oocyte was impaled with two glass microel ctrodes for recording which, when fill d with 3 M

10

15

20

25

KCl, had a tip resistanc of 0.5 to 7.0 megachms. the two electrodes was connected to a differential amplifier via a silver/silver chloride half cell. The bath potential was measured by connecting the other side of the differential amplifier to the bath via a silver/silver chloride pellet and a Ringer/Agar bridge. A low noise, high compliance, voltage-clamp system (NPI) was used to control the membrane potential and to measure membrane current. The oocyte membrane potential was maintained at -60 mV (inside cell negative). One millimolar glutamate (Sigma), 100 μ M quisqualate (Sigma), 1 mM carbamylcholine (Sigma), and the other drugs used in this assay were applied by switching the perfusing medium to a medium containing a drug for approximately three minutes, and the membrane current was recorded on a chart recorder (Linear Instruments).

After impaling the oocyte with the two microelectrodes, and imposing the voltage-clamp, the membrane current (the holding current) gradually declines to a steady state over a period of several minutes. When the holding current stabilizes, so that the chart record is horizontal, the drug is applied for one to three minutes. An oocyte is judged to have a positive response if a rapid inward current spike (downward deflection on the chart), followed by slow current oscillations of decreasing magnitude, is observed. Our lower limit of detection depended on the steadiness of the holding current prior to drug application, but was in the range of 5-10 nA.

30

Construction of pVEGT1

To permit transcription of cloned cDNA without prior endonuclease digestion, bacteriophage T7 transcriptional terminators were added to a cloning v ctor. Plasmid pVEGT' is d scribed in copending U.S.S.N. 07/581,342, which is incorporated by reference h rein. The s qu nce of th putativ T7 RNA transcription terminator, which li s between gene 10 and

15

20

25

30

35

gene 11 of bacteriophage T7, is disclosed by Dunn and Studier (J. Mol. Biol. 166: 477-536 (1983)). As shown in Figure 5, four synthetic oligonucleotides were designed from this sequence and ligated into the vector pGEM-1 (obtained from Promega Biotec, Madison, WI), a plasmid containing a bacterial origin of replication, ampicillin resistance gene, and the T7 promoter adjacent to a multiple cloning site. Terminal phosphates were added to the 5' ends of oligonucleotides ZC776 and ZC777 (Sequence ID Nos. 4 and 5) with T4 polynucleotide kinase and ATP, under standard conditions (Maniatis et al. ibid). sequences of these and other oligonucleotides referred to herein are shown in Table 1.) After the incubation, the kinase was heat killed at 65°C for 10 min. Twenty-five nanograms of oligonucleotide ZC775 (Sequence ID Number 3) and 25 ng of oligonucleotide ZC776 (Sequence ID Number 4) were annealed by incubation at 65°C for 15 minutes, then allowed to cool to room temperature in 500 ml of water. Oligonucleotides ZC777 and ZC778 (Sequence ID Nos. 5 and 6) were similarly annealed. The annealed oligonucleotides were stored at -20°C until use. The vector pGEM-1 was digested with Pst I and Hind III, and the linearized vector DNA was purified by agarose gel electrophoresis. The synthetic T7 terminator (annealed oligonucleotides ZC775, ZC776, ZC777 and ZC778; Sequence ID Nos. 3, 4, 5 and 6) was then cloned into pGEM-1. Twenty-five nanograms of vector plus an equal molar amount of each of the annealed oligonucleotides ZC775/ZC776 (Sequence ID Nos. 3 and 4) and ZC777/ZC778 (Sequence ID Nos. 5 and 6) were combined in a 10 μ l reaction mix. After an overnight ligation at 14°C, the DNA was transformed into competent E. coli JM83 cells, and the transformed cells were selected for ampicillin resistance. Plasmid DNA was prepared from selected transformants by the alkaline lysis procedure (Birnboim and Doly, Nuc. Acids Res. 7:1513-1523 (1979)). A portion of th DNA from these samples was cut with Pst I and Hind III and analyzed on a 4% polyacrylamid gel to identify

clon s that releas d an 80 bp Pst I-Hind III fragment. Other diagnostic cuts, such as Eco RI and Not I, were also made. One of the isolates, designated pGEMT, was shown by restriction analysis to contain the T7 terminator fragment.

Table 1

Oligonucleotide Sequences (5' - 3')

ZC775 (Sequence ID Number 3):

GCT AGC ATA ACC CCT TGG GGC CTC TAA ACG GGT CT

ZC776 (Sequence ID Number 4):

CTC AAG ACC CGT TTA GAG GCC CCA AGG GGT TAT GCT AGC TGC A

2C777 (Sequence ID Number 5):

TGA GGG GTT TTT TGC TGA AAG GAG GAA CTA TGC GGC CGC A

ZC778 (Sequence ID Number 6):

AGC TTG CGG CCG CAT AGT TCC TCC TTT CAG CAA AAA ACC C

ZC1751 (Sequence ID Number 7):

AAT TCT GTG CTC TGT CAA G

25

10

ZC1752 (Sequence ID Number 8):

GAT CCT TGA CAG AGC ACA G

30

ZC2063 (Sequence ID Number 9):

GAT CCA AAC TAG TAA AAG AGC T

35

ZC2064 (Sequence ID Number 10):

CTT TTA CTA GTT TG

(Table 1, continued)

ZC2938 (Sequence ID Number 11):

GAC AGA GCA CAG ATT CAC TAG TGA GCT CTT TTT TTT TTT TTT T

ZC3015 (Sequence ID Number 12):

TTC CAT GGC ACC GTC AAG GCT

10

ZC3016 (Sequence ID Number 13):

15 AGT GAT GGC ATG GAC TGT GGT

ZC3652 (Sequence ID Number 14):

20 ACA TGC ACC ATG CTC TGT GT

ZC3654 (Sequence ID Number 15):

AGT GAT GGC ATG GAC TGT GGT

25

30

35[

(Rosenberg et al., Gene 56:125-135 (1987)) was added to plasmid pGEMT. Plasmid pGEMT was digested with Bam HI and plasmid pAR2529 was digested with Bam HI and Bgl II (Figure 1). The Bam HI-Bgl II terminator fragment from pAR2529 was purified by agarose gel electrophoresis. The terminator fragment was ligated to Bam HI digested pGEMT, and the DNA was transformed into competent E. coli LM1035 cells. Colonies that were ampicillin resistant were inoculated into 5 ml cultures for overnight growth. Plasmid DNA prepared by the alkaline lysis procedure was screened for proper terminator orientation by Bam HI-Sal I dig stion and electrophor sis on an 8% polyacrylamid gel. A clon that contained the terminator in the correct ori ntation, as evidenced by the presenc of a

10

15

25 `

30

35

130 bp Bam HI-sal I fragment, was chos n and named pGEMTT (Figure 1).

To allow pGEMTT to be packaged as single-stranded DNA in the presence of M13 phage proteins, the M13 intergenic region from pUC382 (similar to pUC118 and 119 as disclosed by Vieira and Messing, Methods Enzymol. 153: 3-11 (1987) was added to pGEMTT (Figure 1). Plasmid pGEMTT was digested with Fsp I and Nar I, and the fragment containing the T7 promoter and transcription terminator was purified. Plasmid pUC382 was digested with Fsp I and Nar I, and the fragment encoding the ampicillin resistance gene and the M13 intergenic region was gel purified. These fragments were then ligated together in the presence of T4 DNA ligase. The ligated DNA was transformed into competent E. coli LM1035 cells. Plasmid DNA from twelve ampicillin-resistant colonies was prepared by the alkaline lysis method, and the DNA was screened by digestion with Ava I. The appropriate construction gave two bands, one of 2430 bp and another of 709 bp. One such isolate was chosen and named pVEG. Synthetic oligonucleotides encoding the prime sequence were added to pVEG between the Bam HI and Eco RI sites (Figure 1). Plasmid pVEG was digested with Bam HI and Eco RI and the vector fragment was gel purified. Ninety-six nanograms each of oligonucleotides ZC1751 and ZC1752 (Sequence ID Nos. 7 and 8) were annealed in 4.5 μ l of 10 mM Tris pH 7.5, 20 mM MgCl, and 10 mM NaCl at 65°C for 20 minutes, then the mixture was cooled to room temperature over a period of 30 minutes. The annealed oligonucleotides were ligated to the pVEG vector fragment with T4 DNA ligase and then transformed into competent E. coli LM1035 cells. After growing overnight to develop the colonies, a filter lift was taken of the colonies on the agar plate. The filter was probed with 32P-labeled oligonucl otid ZC1751 (Sequence ID Number 7). All of the colonies were positive. Plasmid DNA was prepared from cultures grown from 12 of th colonis. The plasmid DNA was screened by dig stion with Sst I to verify th

15

20

25.

30

35

absence of the Sst I site between the Eco RI and Bam HI sites of pVEG. All 12 f the plasmid DNAs were negative for Sst I digestion. One of these 12 isolates was chosen and named pVEG'.

A polyadenylate sequence derived from an Aspergillus alcohol dehydrogenase cDNA was added to pVEG. As shown in Figure 1, plasmid pM098 (disclosed in published European patent application EP 272,277 and deposited with American Type Culture Collection under accession number 53428) was digested with Dra I and Bam HI, and the approximately 150 bp poly(A) fragment was purified by agarose gel electrophoresis. This fragment contained mostly poly(A) sequence with very little flanking cDNA. To clone the poly(A) cDNA fragment into pVEG, pVEG was digested with Bam HI and Sma I, and the 3.4 kb vector fragment was gel purified. The vector and poly(A) fragments were ligated together with T4 DNA ligase to produce vector pVEGT (Figure 1).

Synthetic oligonucleotides encoding the prime sequence were added to pVEGT. To accomplish this, pVEGT was digested with Not I and Sst I, and the 370 bp fragment containing the poly(A) sequence and the two T7 transcriptional terminators was purified by agarose gel electrophoresis. Plasmid pVEG' was digested with Not I and Bam HI, and the 3.2 kb vector fragment was gel-purified. Two oligonucleotides (ZC2063 and ZC2064; Sequence ID Nos. 9 and 10) that formed, when annealed, a Bam HI-Sst I adapter were synthesized. The two oligonucleotides were individually kinased and annealed, and ligated with the linearized vector and the poly(A)-terminator fragment. The resultant vector, designated pVEGT' (Figure 1), contained a T7 RNA transcription promoter, an Eco RI cloning site flanked by the prime sequence, a poly(A) tract, and two T7 RNA polymeras terminators.

15

20

25

30

Construction of cDNA library from rat cerebellum poly(A)+
RNA

Because there was evidence suggesting that the GluR was encoded a very large mRNA of 7 kb (Fong, Davidson, and Lester, Synapse 2:657 (1988)) and because full length cDNA encompassing the coding sequence is required for functional cloning of cDNA, measures were taken to optimize for synthesis of large cDNA. A novel method of cDNA synthesis was developed which yielded large full length cDNA. This was evident by demonstration that full length 7.5 kb cDNA could be synthesized from a model 7.5 kb mRNA and that large full length cDNA were present in a library constructed from poly(A) + RNA as demonstrated by Southern blot analysis. In addition, all enzymes which were important in this method were pretested and selected from a large number of lots of enzymes available from commercial suppliers. Once a satisfactory lot was identified, a large amount of the enzyme was purchased and the enzyme was stored at -70°C until used. Once used, the enzyme was stored at -20°C for a few months and then discarded. Different "lots" of enzymes from commercial suppliers, including lots of Superscript reverse transcriptase (BRL), E. coli DNA polymerase I (Amersham) and Mung bean nuclease (NEB), which were used in the cDNA synthesis, were screened for quality in test synthesis assays. Superscript reverse transcriptase lots were assayed for the ability to synthesize unit length (7.5 kb) first strand cDNA from 7.5 kb RNA (BRL) control. Conditions for first strand synthesis with Superscript reverse transcriptase lots were prepared as described below. Radiolabeled first strand cDNA was analyzed by alkaline agarose gel electrophoresis. Superscript lots capable of producing unit length, 7.5 kb cDNA were selected for use.

E. coli DNA polym rase I lots were assayed for the ability to produce, by hairpin DNA formation, full-1 ngth second strand cDNA from the 7.5 kb unit-length first strand cDNA. The second strand cDNA synth ses w re

10

15 ⁻

20

25

30

35

carried out as described below. The quality of the second strand syntheses were assessed by alkalin agarose electrophoresis of the radiolabeled product. DNA polymerase I lots capable of producing 15 kb second strand DNA from the 7.5 kb unit length first strand cDNA were selected for use.

Mung bean nuclease lots were tested for the ability to clip the hairpin DNA formed during second strand synthesis without degrading the cDNA. In addition, varying concentrations of enzyme were added to determine the optimum enzyme concentration for the conditions set forth below. The reactions were assessed by alkaline agarose electrophoresis. Lots and concentrations resulting in the production of 7.5 kb unit length cDNA were selected for use.

Total RNA was prepared from rat cerebella using guanidine isothiocyanate (Chirgwin et al. <u>Biochemistry</u> 18:52-94 1979) and CsCl centrifugation (Gilsin et al. <u>Biochemistry</u> 13:2633-2637 1974). Poly(A)+ RNA was selected from the total RNA using oligo d(T) cellulose chromatography (Aviv and Leder, <u>Proc. Natl. Acad. Sci. USA</u> 69:1408 (1972)).

First strand cDNA was synthesized from one time poly d(T)-selected cerebellum poly(A)+ RNA in two separate reactions. One reaction, containing radiolabeled dATP, was used to assess the quality of first strand synthesis. The second reaction was carried out in the absence of radiolabeled dATP and was used, in part, to assess the quality of second strand synthesis. Superscript reverse transcriptase (BRL) was used specifically as described below. A 2.5x reaction mix was prepared at room temperature by mixing, in order, 10 μ l of 5x reverse transcriptase buffer (BRL; 250 mM Tris-HCl pH 8.3, 375 mM KCl, and 15 mM MgCl₂), 2.5 μ l 200 mM dithiothreitol (made fresh or stored in aliquots at -70°C) and 2.5 µl of a d oxynucl otide triphosphat solution containing 10 mM ach of dATP, dGTP, dTTP and 5-m thyl dCTP (Pharmacia). The reaction mix was

15

20

25

30

35

aliquoted into two tubes of 7.5 μ l each. To the first tube, 1.3 μ l of 10 μ Ci/ μ l α ³²P-dATP (Amersham) was added and 1.3 μ l of water was added to the second reaction Seven microliters from each tube was transferred to reaction tubes. Fourteen microliters of a solution containing 10 μ g of cerebellum poly(A)+ RNA diluted in 14 μ l of 5 mM Tris-HCl pH 7.4, 50 μ M EDTA was mixed with 2 μ l of 1 μ g/ μ l first strand primer, ZC2938 (Table 1; Sequence ID No. 11), and the primer was annealed to the RNA by heating the mixture to 65°C for 4 minutes, followed by chilling in ice water. Eight microliters of the RNA-primer mixture was added to each of the two reaction tubes followed by 5 μ l of 200 U/ μ l Superscript reverse transcriptase (BRL). The reactions were mixed gently, and the tubes were incubated at 45°C for 30 minutes. After incubation, 80 μ l of 10 mM Tris-HCl pH 7.4, 1 mM EDTA was added to each tube, the samples were vortexed and centrifuged briefly. Three microliters of each reaction was removed to determine total counts and TCA precipitable counts (incorporated counts). Two microliters of each sample was analyzed by alkaline gel electrophoresis to assess the quality of first strand synthesis. The remainder of each sample was ethanol precipitated. The nucleic acids were pelleted by centrifugation, washed with 80% ethanol and air dried for ten minutes. The first strand synthesis yielded 1.4 μg of cerebellum cDNA or a 28% conversion of RNA into DNA.

Second strand cDNA synthesis was performed on the RNA-DNA hybrid from the first strand reactions under conditions which encouraged first strand priming of second strand synthesis resulting in DNA hairpin formation. The nucleic acid pellets containing the first strand cDNA were resuspended in 71 μ l of water. To assess the quality of second strand synthesis, α^{32} P-dATP was add d to the unlabeled first strand cDNA. To ncourage formation of the hairpin structure, all reagents xcept th enzymes w re brought to room temperatur, and th reaction mixtures were s t up at

15

20

25

30

35

room temperature. (Alternatively, the reagents can be on ice and the reaction mixture set up at room temperature and allowed to equilibrate at room temperature for a short time prior to incubation at 16°C.) Two reaction tubes were set up for each synthesis. One reaction tube contained the unlabeled first strand cDNA and the other reaction tube contained the radiolabeled first strand To each reaction tube, 20 μ l of 5x second strand buffer (100 mm Tris, pH 7.4, 450 mM KCl, 23 mM MgCl2, 50 mM (NH₄)₂SO₄), 3 μ l of beta-NAD and 1 μ l of a deoxynucleotide triphosphate solution containing 10 mM each of dATP, dGTP, dTTP and dCTP (Pharmacia), 1 μ l α^{32} P-dATP or 1 μ l of water (the radiolabeled dATP was added to the tube containing the unlabeled first strand CDNA), 0.6 μ l of 7 U/μ l E. coli DNA ligase (Boehringer-Mannheim), 3.1 μ l of 8 U/ μ l E. coli DNA polymerase I (Amersham), and 1 μ l of 2 U/ μ l of RNase H The reactions were incubated at 16°C for 2 hours. (BRL). After incubation, 3 μ l was taken from each reaction tube to determine total and TCA precipitable counts. Two microliters of each sample was analyzed by alkaline gel electrophoresis to assess the quality of second strand synthesis by the presence of a band of approximately twice unit length. To the remainder of each sample, 2 μ l of 2.5 μ g/ μ l oyster glycogen, 5 μ l of 0.5 M EDTA and 200 μ l of 10 mM Tris-HCl pH 7.4, 1 mM EDTA were added, the samples were phenol-chloroform extracted, and isopropanol precipitated. The nucleic acids were pelleted by centrifugation, washed with 80% ethanol and air dried. The yield of double stranded cDNA in each of the reactions was approximately 2 μ g.

The single-stranded DNA in the hairpin structure was clipped using mung bean nuclease. Each second strand DNA sample was resuspended in 12 μ l of water. Two microliters of 10x mung bean buffer (0.3 M NaOAC, pH 4.6, 3 M NaCl, 10 mM ZnSO₄), 2 μ l of 10 mM dithiothreitol, 2 μ l of 50% glyc r l, and 2 μ l of 10 U/ μ l mung b an nuclease (NEB, lot 7) were add d to each tub , and the reactions

were incubated at 30°C for 30 minutes. Aft r incubation, 80 μ l of 10 mM Tris-HCl pH 7.4, 1 mM EDTA was added to each tube, and 2 μ l of each sample was subjected to alkaline gel electrophoresis to assess the cleavage of the second strand product into unit length cDNA. One hundred microliters of 1 M Tris-HCl pH 7.4 was added to each sample, and the samples were twice extracted with phenol-chloroform. Following the final phenol-chloroform extraction, the DNA was isopropanol precipitated. The DNA was pelleted by centrifugation, washed with 80% ethanol and air dried. Approximately 2 μ g of DNA was obtained from each reaction.

The cDNA was blunt-ended with T4 DNA polymerase after the cDNA pellets were resuspended in 12 μ l of water. Two microliters of 10x T4 DNA polymerase buffer (330 mM Tris-acetate, pH 7.9, 670 mM KAc, 100 mM MgAc, 1 mg/ml gelatin), 2 μ l of 1 mM dNTP, 2 μ l 50 mM dithiothreitol, and 2 μ l of 1 U/ μ l T4 DNA polymerase (Boehringer-Mannheim) were added to each tube. After an incubation at 15°C for 1 hour, 180 μ l of 10 mM Tris-HCl pH 7.4, 1 mM EDTA was added to each sample, and the samples were phenol-chloroform extracted followed by isopropanol precipitation. The cDNA was pelleted by centrifugation, washed with 80% ethanol and air dried. Eco RI adapters (Invitrogen, Cat. # N409-20) were ligated to the blunted cDNA after the DNA from each reaction was resuspended in 6.5 μ l water.

The first strand primer encoded an Sst I cloning site to allow the cDNA to be directionally cloned into an expression vector. The cDNA was digested with Sst I followed by phenol-chloroform extraction and isopropanol precipitation. After digestion, the cDNA was electrophoresed in a 0.8% low melt agarose gel, and the cDNA over 4.2 kb was electroeluted using an Elutrap (Schleicher and Schuell, Keene, NH). The electroeluted cDNA in 500 μ l of buff r was isopropanol precipitat d and the cDNA was pelleted by centrifugation. The cDNA pellet was washed with 80% ethanol.

10

15

20

25

30

35

A cerebellum cDNA library was established by ligating the cDNA to the Eco RI-Sst I digested, agarose gel purified pVEGT'.

Ten sublibraries of one million clones each were constructed representing a library of ten million independent clones. To prepare each sublibrary, 80 ng of linearized vector were ligated to 40 ng of cDNA. After incubation at room temperature for 11 hours, 2.5 μg of oyster glycogen and 80 μ l of 10 mM Tris-HCl, 1 mM EDTA was added and the sample was phenol-chloroform extracted followed by ethanol precipitation. The DNA was pelleted by centrifugation, and the DNA pellet washed with 80% ethanol. After air drying, the DNA was resuspended in 3 µl of water. Thirty-seven microliters of electroporation-competent DH10B cells (BRL) was added to the DNA and electroporation was completed using a BioRad electroporation unit. After electroporation, 4 ml of SOC (Maniatis et al.) was added to the cells, and 400 μ l was spread on each of 10-150 mm LB ampicillin plates. Each plate represented a sublibrary of 100,000 clones. After an overnight incubation, the cells were harvested by adding 10 ml of LB ampicillin media to each plate and scraping the cells into the media. Glycerol stocks and plasmid DNA were prepared from each plate. The library background (vector without insert) was established at about 15%.

Detection of Glu R activity from the cDNA library

The <u>Xenopus</u> oocyte efficiently translates exogenously added mRNA. Preliminary experiments were done using the mouse ml muscarinic receptor cDNA (a G protein-coupled receptor that can be detected by voltage-clamp) cloned into pVEGT'. Injection of RNA transcribed in <u>vitro</u> from increasing dilutions of the ml template DNA indicated that ml agonist induced activity could b d t cted for one clone in a pool size f 100,000. A cerebellum sublibrary was plated into ten pools of 100,000 unique clones.

10

15

20

25

30

Th pools could also b replica plated onto a nitrocellulose filter and th original and replica allowed to grow for a few hours. The original plate is scraped to harvest all the colonies. Plasmid DNA is prepared and purified by cesium chloride gradient ultracentrifugation. The DNA from each pool is transcribed in vitro with T7 RNA polymerase in the presence of 7-methyl-G, the capped nucleotide, to increase translation efficiency. Template DNA transcription reactions are spiked with a dilution of two control genes cloned into pVEGT': the mouse ml gene and a secreted version of the human placental alkaline phosphatase gene (SEAP; Tate et al., Fed. Am. Soc. EXD. Biol. 8: 227-231 (1990), incorporated by reference herein). Transcription from the control genes would allow selection of oocytes that more efficiently translate the injected RNA, and a determination whether cocytes that are negative for the GlugR are true negatives, that is, still having a detectable ml agonistinduced response.

Plasmid DNA prepared from each of the 10 pools of 100,000 clones, which in total represented one sublibrary of one million clones of the cerebellum cDNA library, was purified by cesium chloride gradient ultracentrifugation. The DNA was transcribed in vitro with T7 RNA polymerase (Pharmacia) in the presence of capped nucleotide (GpppG, Pharamcia). The presence of a poly(A) sequence and two T7 RNA polymerase terminators in pVEGT' resulted in RNA with a capped 5' end, the sequence of the cDNA insert, and 3' poly(A) tails. Capped RNA is believed necessary for efficient translation in oocytes (Noma et al. Nature 319:640 (1986)) and the poly(A) sequence has been shown to increase the synthesis of a protein in oocytes by more then 40 fold. The transcription reaction tubes were set up by adding 12 μ l of 5x transcription buffer (Stratagene Cloning Systems, La Jolla, CA), 3 μ l each of 10 mM ATP, CTP, GTP, and UTP, 6 μ l of 10 mM GpppG (Pharmacia), 6 μ l of 1 mg/ml BSA, 3 μ l of 200 mM DTT, 1.5 μ l of 40 U/ μ l

15

20

25

30

35

RNasin (ProMega Biotech, Madison, WI), 8.5 µl of water, 10 μ l of cDNA containing 5 to 10 μ g DNA, and 1 μ l of 70 $U/\mu l$ T7 RNA polymerase. After mixing, 10 μl of the reaction was transferred to a tube containing 0.5 μ Ci of α^{32} P-UTP to determine the total counts and counts incorporated into RNA. The samples were incubated at 37°C for one hour. The cDNA in the unlabeled samples was degraded with the addition of 1 μ l of 200 mM DTT, 2 μ l of 30 U/ μ l DNase I, and 0.5 μ l of 40 U/ μ l RNasin and the incubation was continued at 37°C for 15 minutes. Forty microliters of water was added to the radiolabeled reactions, and 1 μ l was removed from each sample and counted to determine total counts. The remainder of the labeled samples were ethanol precipitated. The samples were centrifuged to collect the RNA and the RNA pellets were counted to determine the counts incorporated into RNA. After the DNA degradation reaction in the unlabeled samples, 70 µl of 10 mM Tris-HCl, 1 mM EDTA was added to each sample, and the samples were twice-extracted with phenol-chloroform followed by one chloroform extraction. The RNA was ethanol precipitated. After centrifugation to collect the RNA, the pellets were washed with 80% ethanol, followed by air drying for 10 minutes. A typical yield of the unlabeled RNA was 20 to 30 μ g. unlabeled RNA was resuspended at 2 μ g/ μ l in diethylpyrocarbonate (DEPC, Sigma) treated water and stored at -70°C.

Prior to microinjection into oocytes, the RNA samples were thawed and centrifuged in a microfuge for 5 minutes to remove any particles that might clog a microinjection pipet. After centrifugation, 80% of each sample was removed and split into two tubes.

The RNA from each of the 10 sublibraries were injected into oocytes as described above and translation was allow d for four days. Expression of GlugR activity was ass ss d by voltage-clamp assay as described ab ve. One of the 10 sublibrari s, Z93-1.9, pr duced a signal with administration of quisqualate t the oocyte.

<u>subdivision of the cDNA library pool to obtain pure Gluck</u> clone

The DNA pool (Z93-1.9) was subdivided by plating clones from the glycerol stock onto LB ampicillin plates. To determine the number of clones that should be plated for the subdivision of the 100,000 clone pool to identify a positive clone, the probability equation N = ln (1- P) / ln (1 - f) (Maniatis et al., ibid.) was used, where P is the desired probability of including the clone of interest, f is the fraction of positive clones in the pool, and N is the number of clones to be plated to provide the given probability. For a probability of 99.8% for a pool size of 100,000 to contain one positive clone, 621,461 clones should be plated.

Forty-eight 150 mm LB ampicillin plates were plated with the glycerol stock representing the 100,000 positive pool, Z93-1.9, at a density of approximately 14,000 clones per plate to give a total of 670,000 clones. After an overnight incubation 37°C, the bacteria on each plate were harvested into 10 ml of Solution I (as described by Birnboim and Doly, Nuc. Acids Res. 7:1513 (1979)), incorporated by reference herein). A glycerol stock was prepared from a portion of the cells, and plasmid DNA was prepared from the remainder of the cells. Six pools of DNA representing eight of the LB ampicillin plates each were prepared by combining one tenth of the plasmid DNA from groups of eight plates into each pool. The plasmid DNA from these six pools was purified by cesium chloride gradient centrifugation. The DNA was transcribed into RNA as outlined above. Transcription of the parent pool Z95-1.9 was included as the positive control. Oocytes were injected with the RNA and voltage-clamp assays on the oocytes identified pool Z99-25-32 as positive for GlugR. Pool Z99-25-32 contained DNA pr pared from plates 25 through 32.

Plasmid DNA from plates 25 t 32 wer cesium chloride banded and transcribed into RNA as described above al ng with th positive par nt pool Z99-25-32.

.

15

10

·5

20

25

30

35

WO 92/10583 PCT/US91/09422

41

Occyt s were injected with the RNA and voltag clamp assays, carried out as described above, identifi d pools Z104-25 and Z111-32 as being weakly positive, Z106-27 and Z109-30 as intermediately positive, and Z108-29 and Z110-31 as the most positive. The pool resulting in Z110-31 was chosen for further subdivision.

5 .

10

15

20

25

30 ...

35

Identification of positive pools from the subdivision of the positive pool of 14,000 (Z110-31) from the glycerol stock was unsuccessful. Therefore, plasmid DNA prepared from the pool resulting in Z110-31 was electroporated into bacteria and plated on 60 plates at a density of 1,000 clones/plate. Plasmid DNA was prepared from the bacteria harvested from each plate. Aliquots of the plasmid DNA from each plate were mixed to make six pools representing ten plates each. The plasmid DNA was cesium chloride banded, and the RNA was transcribed as described above. RNA was transcribed from pools Z108-29, Z110-31, and a muscarinic receptor cDNA, ml, for use as positive controls. The RNA was injected into occytes and voltage-clamp assays were carried out as described above. The assays identified pool Z133-21 to 30 as positive.

Plasmid DNA from plates 21 to 30 were cesium chloride banded and transcribed as described above. The transcribed RNA and the RNA from the parent pool Z133-21 to 30 were injected into oocytes and assayed as described above. The voltage-clamp assay identified pool Z142-22 as positive.

Identification of positive pools by the subdivision of the positive pool Z142-22 from a glycerol stock proved unsuccessful. Restriction analysis of plasmid DNA prepared from randomly selected clones from pools Z110-31 (the pool of 14,000) and Z142-22 (the pool of 1,000) indicated that 50% of pool Z110 - 31 and 68% of pool Z142 - 22 were clones without inserts.

T ass ss physical methods for enriching for the GlugR clone and to establish how many clones from pool Z142-22 n ded to be assay d to include a GlugR clone, undigested plasmid DNA from p ol Z142-22 was

10

15

20

25

30

35

1 ctrophoresed on an agarose gel. The super-coil band representing v ctor without insert was cut out and the remainder of the DNA was eluted from the gel. The DNA was then electroporated into bacteria cells, and plated at densities of 3,400, 6,900, and 13,800 clones per plate. The plates were replica plated and grown overnight. Plasmid DNA was prepared from the cells harvested from the replica of each plate. The plasmid DNA was transcribed, and the RNA was assayed in oocytes as described above. As a control, each pool contained the equivalent of one colony of ml as an internal positive control. In addition, ml was used as an external positive control. The voltage-clamp assays identified the DNA from the 6,900 clone pool (Z167-7) as positive.

The clones represented on the 6,900 clone plate that resulted in the positive pool Z167-7 were subdivided by replica plating the master plate onto a Biodyne-A nylon membrane on an LB ampicillin plate. The replica plate was incubated four hours at 37°C. After incubation, sub-pools were prepared by removing the membrane from the plate, taping the membrane to a sterile glass plate on a light box, and overlaying the membrane with a grid which divided the membrane into 100 sections. The sections of the grid and underlying membrane were then cut out with a razor blade that had been dipped in alcohol and flamed between each cut. Alcohol-treated, flamed forceps were used to transfer each membrane section to a test tube containing 12.5 ml of LB ampicillin media. The cultures containing the membrane sections were incubated overnight at 37°C. After incubation, 0.5 ml of each culture was mixed with 0.5 ml of 50% glycerol and stored at -70°C to establish glycerol stocks of each sub-pool. Aliquots of the 100 cultures were pooled in a 10 X 10 matrix with samples (1) through (10) on the abscissa and sampl s (a) through (j) n the ordinate. For example, 1 ml of cultur s (1) through (10) were added to tube 1 and 1 ml of cultur s (1), (11),

10

15

20

25

30

35

(21), (31), (41), (51), (61), (71), (81), and (91) w re added to tube (a) and so on until 10 rows f 10 and 10 columns containing pools of 10 cultures each were completed. Ten microliters of an overnight culture containing m1-transformed bacteria was added to each pool as an internal control. Plasmid DNA was prepared from the 20 sub-pools, and the DNA was purified by cesium chloride gradient centrifugation. RNA was transcribed from the plasmid DNA and was assayed in oocytes as described above. Positive controls were the parent pool Z167-7 and pure m1 RNA. The voltage-clamp assays indicated that only pools Z175-1 and Z191-g were positive. Consulting the matrix, this indicated that the membrane section number (7) contained the Glu_GR clone.

To subdivide the clones contained in section (7), a piece of Biodyne A membrane was applied to the master plate containing section (7), the membrane extending beyond section (7) on each side by half the width of section (7). The membrane was removed from the plate, applied to a fresh LB ampicillin plate colony side up, and incubated overnight at 37°C. The membrane was subdivided as described above with the central region of the membrane, the actual section (7) area, divided into 9 small, equivalent-sized squares and the membrane on each side of section (7) was taken as four additional areas. Each membrane section was used to inoculate a 10 ml liquid culture. Bacteria transformed with the ml clone were used as an internal control in each culture as described above. After overnight incubation at 37°C, plasmid DNA was prepared, and the DNA was purified by cesium chloride gradient centrifugation. RNA was transcribed and assayed in oocytes as described above using RNA from ml and the parent pool number (7) as positive controls. GlugR activity was found in only pool Z203-7 corresponding to membrane section number (7).

Pool Z203-7 was subdivided by electroporating the plasmid DNA prepared from the membrane section numb r (7) into DH10B electroporation-competent cells. The

15

25

30

transformants were plat d at a density enabling individual colonies to be pick d. Individuals clon s were picked to a master plate and into 2 ml of LB ampicillin media. The cultures were incubated overnight, and plasmid DNA was prepared by the method essentially described by Holms and Quigley (Anal. Bioc. 114: 193, (1981)). Restriction analysis suggested that the clones were grouped into 7 different classes of clones. Plasmid DNA, prepared from each class, representing fifty total clones were prepared, transcribed, and assayed in oocytes as described above. However, none of the clones were positive.

To screen for positive clones, electroporation-competent E. coli DH10B cells were electroporated with the DNA prepared from membrane section number (7) (Z203-7) and were plated at 180, 360, 900, and 1800 colonies per plate. The plates were incubated overnight, and replica plates were prepared as described above. Plasmid DNA prepared from each replica plate was combined with 1 to 1000 parts of ml as an internal control. The DNA pools, the ml clone and the parent pool Z203-7 were transcribed, and the RNA was assayed by occyte injection. The first transcription and injection showed no positives, however, upon retranscription and reanalysis the 1800 clone pool (Z264-1800) was positive for Gluck activity.

To subdivide the positive pool of 1800 (Z264-1800), all of the colonies from the plate of 1800, 1528 in total, were each picked to two 100 mm LB ampicillin agar plates on a 100 colony grid. After overnight growth, one set of the duplicate plates was designated as a master set and was placed at 4°C. The other set was replica plated to a third set of plates. After overnight incubation of these plates, the cells on the replica plat s were harv sted into media and plasmid DNA was prepared from the pooled cells. As d scribed above, an internal ml control was included in each DNA pr paration. ml DNA and the parent Z264-1800 DNA were

15

20

25

30

35

used as external positive controls. Plasmid DNA pr par d from the 16 plates was transcribed, and the RNA was assayed in oocytes as described above. One of the pools of 100 clones, Z256-I produced Gluck activity.

To identify which clone of the 100 clones from Z256-I produced the Glu_cR activity, a 10 x 10 matrix of the clones was constructed. A liquid culture of each clone was grown. One milliliter of each culture was added to each of two tubes representing the appropriate row and column of the 10 x 10 matrix. As described previously, plasmid DNA encoding m1 was used as an internal positive control. Plasmid DNA prepared from each tube, m1 DNA and DNA from the parent pool Z264-1800 were transcribed and assayed in oocytes as described above. Glu_cR activity was identified only in row (5) and column (e). Thus, the positive clone number 45 was identified as containing the Glu_cR activity.

To confirm the result, plasmid DNA from clone #45 was prepared, transcribed and assayed in oocytes as described above. The results of the assay indicated that clone #45 was capable of producing Gluck activity. Figure 2 illustrates the data taken from voltage-clamp recordings at several stages in the subfractionation of the cerebellum library. Panel (a) is a recorded response to quisqualate of an oocyte previously injected with in vitro transcribed RNA from a rat cerebellum sublibrary of 100,000 independent colonies; panel (b) shows the response to quisqualate in a cell previously injected with RNA transcribed from a subfractionated pool of 14,000 colonies. The peak current was truncated by the chart recorder, but the actual peak current (estimated from a digital panel meter) was approximately 1300 nA. Panel (c) shows the response to quisqualate in a cell injected with pure Gluck RNA from clone 45-A. The amount of RNA inject d per occyte was approximately 100 ng, except in pan 1 (c) wher the amount f RNA was 50 pg.

The following describ s an alternative m ans f r subdividing and screening a positive po 1. W rking with

15

20

25

30

cDNA ins rts in a plasmid based rath r than a lambdabased vector influences the subfractionation protocol. Once a positive pool is identified, the replica filter is overlayed with another sterile nitrocellulose filter. The filter is cut into 88 pieces by using evenly spaced cuts of 10 rows and 10 columns to form a grid. Each of the 88 pieces is transferred to 10 ml of sterile LB +Amp and grown for several hours. Twenty pools are formed; C 1-10 (corresponding to column number) and R 1-10 (corresponding to row number). An aliquot of each of the 88 subfractions is pipetted into 2 tubes, corresponding to its position in a row and a column. DNA is isolated from the 20 pools, purified on CsCl gradients and transcribed in an in vitro reaction that includes the control m1 and SEAP plasmids. After injection into oocytes and voltage-clamp recording there are 2 positive pools, pinpointing the location of 1 of the 88 original subfractions.

Because the positive clone is still part of a pool it must be further subdivided. The probability equation described above is used to determine the number of clones to be plated for the next subdivision of the pool. The glycerol stock from the positive pool is plated out at, e.g., 3000, 6000 and 18,000 clones per plate. After replica plating the DNA is harvested, transcribed, injected and assayed. The pool which is positive is subdivided into a grid of 88 as described above. assay is repeated, and a single square of the grid is positive. At the next step of subdivision of the pool, 100 individual colonies to a plate are picked, replica plated, and 20 pools are made for transcription and assay. Positive clones are streaked out, several colonies picked and restriction mapped and template and transcript prepared for injection and assay.

35

10

15

20

25

30

35

Characterization of Glu R

To establish that th Gluck encod d by clone 45-A couples to G-protein, clone 45-A Gluck RNA was transcribed and injected into occytes as described above. Two days after injection the oocytes were divided into control and toxin-treated groups. The oocytes in the toxin-treated group were treated with a final concentration of 4 μ g/ml of B. pertussis toxin (List Biological Laboratories Inc., Campbell, CA), and both groups were incubated for 24 hours at 19°C as described by Sugiyama et al., Nature 325:531 (1987) and Moriarity et al., J. Biol. Chem. 264:13521 (1989), both of which are incorporated by reference herein. The oocytes from both the control and toxin-treated groups were subjected to voltage-clamp assays as described previously. In one example, oocytes perfused as described previously with 100 μM L-glutamic acid showed a mean L-glutamic acid-induced current of 264.2 nA +/- 73 nA in control oocytes (SEM, n=6) and 57.7 nA +/- 19 nA (n=9) in toxin-treated oocytes. The mean membrane current in the toxin-treated group was significantly smaller (p < 0.01) than in the control group suggesting that oocytes injected with 45-A RNA coupled to a pertussis toxin-sensitive G protein.

L-glutamic acid and some of its structural derivatives that are known to activate Gluck currents in a dose-dependent manner were applied to oocytes that had been injected with RNA transcribed from the 45-A clone. RNA was transcribed and oocytes were prepared and injected as previously described. Dose dependent responses were measured using voltage clamp assays were carried out in the presence of increasing concentrations of L-glutamic acid (Sigma), quisqualic acid (Sigma), ibotenic acid (Sigma), or trans 1-amino-cyclopentyl-1,3 dicarboxylic acid (tACPD; Tocris Neuramin, Essex, England). Four or five separate ocytes were perfused with increasing concentrations of a particular drug with 30 minutes betw en consecutive applications of the drug to minimize any int rf rence from desensitization. The

10

15

20

25

30

35

responses were normalized to a subsequent response to 100 μ M L-glutamic acid. The data were analyzed using the following equation:

(Fractional current) = (Doseⁿ)/(Doseⁿ) + (EC₅₀)ⁿ, where:

Dose = a dose of drug normalized to that evoked by a subsequent application of 100 μM L-glutamic acid;

Fractional current = the peak current evoked by a dose, as defined above;

 EC_{50} = effective concentration that evokes a 50% response (a measure of the potency of an agonist); and

n = the Hill coefficient, a measure of the cooperativity of the reaction.

Using this equation, the effective concentration at 50% stimulation relative to 100 μ M L-glutamic acid was determined for each dose response experiment. Figure 6 shows a representative dose response curve for varying concentrations of L-glutamic acid. The potency series of glutamate analogs and their associated EC₅₀'s are listed in Table 2.

Table 2

Glutamate Analog Potencies (EC₅₀)

Quisqualic acid 0.681 μ M L-glutamic acid 12.32 μ M Ibotenic acid 32.37 μ M tacpD 376 μ M

In addition, oocytes were exposed to the following L-glutamic acid analogs: aspartic acid (Tocris Neuramin), kainic acid, N-methyl-D-aspartic acid (NMDA; Sigma), 2-amino-4-phosphonobytyric acid (APB; Sigma), c-amino-3-hydroxy-5-methyl-isoxazole-4-propionic acid (AMPA; Research Biochemicals Inc., Wayland, MA) at saturating conc ntrations and the responses were each negligible to a subsequent response to 100 μ M L-glutamate. The L-glutamic acid analogs that were found to be ineffective were 1 mM aspartic acid, 1 mM kainic

15

20

25

30

35

acid, 100 μ M NMDA + 10 μ M glycine, 100 μ M APB and 100 μ M AMPA.

Voltage clamp assays were also carried out on injected oocytes to measure the inhibition by the putative glutamate G protein-coupled receptor antagonist, 2-amino-3-phosphonopropionic acid (AP3). Voltage clamp assays showed that at 1 mM, DL-AP3 (Sigma) reduced the current evoked by 10 μ M glutamic acid to 59.3 +/- 7.3% of the control.

Clone 45 cells were streaked out on LB Amp plates and several colonies were picked, grown up and the DNA isolated. Pure 45-A DNA was prepared and restriction mapped by standard procedures. Clone 45-A has been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD, 20852, under ATCC Accession No. 68497. DNA was digested with single or multiple enzymes. The fragments were separated on both 1% agarose and 4% Nusieve gels by electrophoresis. After electrophoresis the DNA was transferred to nitrocellulose filters using standard protocols for Southern transfer. Restriction sites were mapped based on size and based on hybridization to Pst I subclones of 45-A DNA. Additionally, the entire 45-A cDNA insert can be isolated by digestion with Not I restriction endonuclease. The Not I insert was kinased with $\gamma^{-32}P$ ATP, and after digestion of half of the sample with Bam HI to remove the 3' label, both samples were subjected to digestion with a number of enzymes known to be present once in the insert. In this way the unique sites could be localized. A restriction map of GlugR clone 45-A is shown in Figure 3.

The entire 45-A clone was sequenced in both directions using the dideoxynucleotide chain termination method (Sanger and Coulson, J. Mol. Biol. 94:441 (1975), incorporated herein by reference). Figure 5 (Sequence ID Nos. 1 and 2) shows the DNA sequence and deduced amino acid sequence of clone 45-A. Figure 5 also shows the location of putative N-linked glyc sylation sites, which

10

15 ·

20

25

hav been predicted to occur at th amino acid s quence Asn-X-Thr.

As shown in Figure 5, seven putative transmembrane domains have been predicted from the deduced amino acid sequence of clone 45-A using the method described by Eisenberg et al. J. Mol. Biol. 179:125-142, (1984), incorporated herein by reference. Only those predicted to be transmembrane multimeric domains were included. An additional transmembrane domain (the third) was predicted using the method of Hopp and Woods, Proc. Natl. Acad. Sci. USA 78:3824-3838 (1981). Based on these predictions, the protein encoded by clone 45-A appears to have two unusually large domains on the amino- and carboxy-termini that are not found in any of the other reported G protein-coupled receptors which have the common structural feature of seven predicted membrane spanning regions. Analysis of the deduced amino acid sequence of clone 45-A predicts three other hydrophobic stretches including one at the amino-terminus of the sequence. This amino-terminal hydrophobic stretch may be a signal sequence, although no signal cleavage site is predicted downstream of the sequence.

Poly(A)+ RNA was isolated from total rat brain and rat cerebellum using oligo d(T) cellulose chromatography as described by Aviv and Leder (ibid.). Poly(A)+ RNA from rat retina, rat heart, rat lung, rat liver, rat kidney, rat spleen, rat testis, rat ovary and rat pancreas were purchased from Clonetech. The poly(A)+ RNA samples were analyzed by northern analysis (Thomas, Proc. Natl. Acad. Sci. USA 77:5201-5205 (1980), which is incorporated by reference herein). The RNA was denatured in glyoxal, electrophoresed in agarose and transferred to a nitrocellulose membrane essentially as described by Thomas (ibid.). The northern blot was hybridized with a radiolabeled 3473 bp Ec RI-Xba I fragm nt from th 45-A clon. Autoradiography of th blot show d hybridization to a maj r band f approximat ly 7 kb and a smaller band

15

20

25

30

35

of approximately 3.8 kb in the total rat brain and rat cerebellum RNA.

Single-stranded cDNA was synthesized using 1 μ g of the poly(A) + RNA using Superscript reverse transcriptase (BRL) under conditions described by the manufacturer. One fourth of the cDNA was used as a template for PCR amplification using 40 pmoles each of the GluGR-specific primers ZC3652 (Table 1; Sequence ID Number 14) and ZC3654 (Table 1; Sequence ID Number 15) and 2.5 U Taq I polymerase (Perkin Elmer Cetus, Norwalk, VA) and conditions specified by the manufacturer. As an internal control, the PCR reaction also contained 2 pmoles each of the glucose-6-phosphate dehydrogenase-specific primers ZC3015 (Table 1; Sequence ID Number 12) and ZC3016 (Table 1; Sequence ID Number 13). After thirty cycles (one minute at 94°C, one minute at 60°C, ninety seconds at 72°C), the samples were phenol-chloroform extracted and 20% of each reaction was electrophoresed in agarose. The DNA was bidirectionally transferred to nitrocellulose membranes, and the filters were hybridized with either radiolabeled ZC3652, ZC3654, ZC3015 and ZC3016 (Sequence ID Nos. 14, 15, 12 and 13, respectively) or with the radiolabeled Eco RI-Xba I fragment of clone 45-A described above. Autoradiography of the hybridized blot showed that Gluck transcript was mainly confined to total rat brain and rat cerebellum; however, longer exposures showed a Glu_R-specific transcript in both retina and testis.

Total RNA was prepared, as described above, from specific rat brain regions including frontal cortex, cerebellum, hippocampus, cortex, striatum, pons medulla, and the remainder of the brain. Single-stranded cDNA was synthesized as described previously using 20 μ g of total RNA in 50 μ l using Superscript reverse transcriptase (BRL) under conditions described by th manufactur r. After a one hour incubation at 42°C, the samples w re treated with RNAse (Boehringer Mannheim Biochemicals, Indianapolis, IN), ph nol-chloroform extracted, and

10

15

20

25

30

35

ethanol precipitated. The samples were resusp nded in water and half of each sample was subj cted to PCR amplification. Each PCR amplification contained 40 pmoles of each of the GlugR-specific primers ZC3652 and ZC3654 described above (Sequence ID Numbers 14 and 15), 2 pmoles of each of the glucose-6-phosphate dehydrogenase-specific primers ZC3015 and ZC3016 (Sequence ID Nos. 12 and 13) and 2.5 U Taq I polymerase (Perkin Elmer Cetus) and conditions described by the manufacturer. After 35 cycles (one minute at 94°C, one minute at 60°C, ninety seconds at 72°C), the samples were phenol-chloroform extracted, and 20% of each reaction was electrophoresed in agarose. The DNA was transferred to a nitrocellulose membrane, and the filter was hybridized with the radiolabeled Eco RI-Xba I fragment of clone 45-A described above. Autoradiography of the hybridized blots showed a broad distribution of the GlugR transcript throughout the brain, although the frontal cortex and cerebellum appear to be somewhat enriched.

Southern analysis of rat and human genomic DNA was carried out using the method essentially described by Blin et al. (Nuc. Acids Res. 3:2303 (1976), which is incorporated by reference herein). Briefly, rat and human genomic DNA was prepared from the rat cell line UMR 106 (ATCC CRL 1661) and a human hepatoma cell line (ATCC HTB 52), respectively. The genomic DNA was digested with either Eco RI or Pst I, and electrophoresed through agarose. The DNA was transferred to a nitrocellulose membrane, and the membrane was hybridized with a radiolabeled 1.6 kb Pst I fragment from clone 45-A. Autoradiography of the hybridized blot suggest that the human gene has a similar sequence to the rat Gluck sequence, the Gluck gene contains at least one intron, and that there are a small number of closely related genes.

Expression in Mammalian C lls

The entire GlugR cDNA insert was removed from the pVEGT' cloning v ctor by digestion with Not I and Xba I.

15

20

25

30

35

The ends wer blunt d with DNA polymerase I (Klenow fragm nt) and dNTPs, and were then ligated with Eco RI (Smart) linkers. After linker ligation, the insert with Eco RI ends in kinased and ligated to Eco RI-cut and capped Zem228 expression vector. Bacteria were transformed with the ligation reaction and clones were characterized by restriction analysis and partial sequencing (see Fig. 4).

Cultured mammalian cells, such as BHK 570 and BHK ts13 served as host cells for expression. Twenty five μg of CsCl-purified DNA was precipitated with calcium phosphate and added to tissue culture cells in a 150 mm plate. After 4 hours the cells were subjected to a glycerol shock and were then put into non-selective medium. In some cases it may be necessary to include an antagonist to the GlugR in the medium to prevent expression of a cytotoxic response in those cells where the Gluck is expressed at levels high enough to cause a certain amount of autoactivation. Transiently expressed Gluck ligand binding activity or PLC activation, cells are harvested after 48 hours. Stable expression was detected after 2 weeks of selection. The Zem228 expression vector includes a promoter capable of directing the transcription of the Glu R gene, and a selectable marker Resistance for the bacterial neomycin resistance gene. to the drug G-418, an inhibitor of protein synthesis, was used to identify stably transfected clones. Presence of the SV40 ori region on the vector allows the expression construction to also be used for transient expression. In some instances it was preferable to include DNA for another selectable marker, the DHFR gene, in the transfection protocol. Selection with both G-418 and methotrexate allowed isolation of clones whose expression of Gluck can be subsequently amplified by the addition of incr asingly higher concentrations of methotrexate to the culture medium.

Transfected c 11 lines expressing Glu_GR wer identified by the binding of 3H -glutamate t membrane

15

20

25

preparations from transf cted cells. Cell lines expr ssing low to moderate levels of GlugR ar used to set up functional screening assays.

Clones of BHK 570 and BHK TK ts13 cells expressing the rat G protein-coupled glutamate receptor cDNA were plated in two or three 150 mm maxi-plates culture dishes and were grown to confluency. The cells from each plate were scraped in 5 ml of PBS (phosphate buffered saline, Sigma Chemical Co., St. Louis, MO), which was was pre-chilled to 4°C. The cells were removed to a chilled centrifuged tube, and the plates were each rinsed with 5 ml of chilled PBS and pooled with the cells. The chilled tubes were spun at 1,000 rpm for two minutes, and the supernatant was discarded. The cells were frozen at either -70°C or on dry ice. In some cases, the cells were left overnight at -70°C. The cells were thawed on ice and were resuspended in 10 ml of a buffer containing 30 mM Tris, pH 7.0, 2.5 mM CaCl2, 1 mM PMSF, which was pre-chilled to 4°C, by homogenizing the cells for about 15 seconds. The suspension was poured into chilled centrifuge tubes. The homogenizer was rinsed with 10 ml of the same chilled solution, and the rinse was combined with the suspension. The centrifuge tubes were spun for fifteen minutes at 40,000 x g at 4°C, and the supernatant was discarded. The pellet was homogenized with a buffer containing 30 mM Tris, pH 7.0, 2.5 mM CaCl2, which was pre-chilled to 4°C. The homogenizer was rinsed with the chilled buffer, and the rinse was combined with the homogenate. The homogenate was spun as described above. The second homogenization was repeated on the resulting pellet. The final pellet was resuspended in between two and five milliliters of 30 mM Tris, pH 7.0, 2.5 mM CaCl, which was pre-chilled to 4°C. Triplicate samples were prepared for each plus and minus quisqualate assay point such that 250 μ l aliquots of each homog nate sample w r add d to th w lls of a 96-well microtiter plate. To a buff r containing 30 mM Tris, pH 7.0, 2.5 mM CaCl2, which was pre-chilled to 4°C, a final concentration of 10 nM

15

20

25

30

35

tritiated glutamic acid was added, and the solution was split in half. To one half, quisqualate was added to a final concentration of 1 mM. Two hundred and fifty microliter aliquots of either 30 mM Tris, pH 7.0, 2.5 mM CaCl₂, 5 nM tritiated glutamic acid and 500 mM quisqualate, or 30 mM Tris, pH 7.0, 2.5 mM CaCl₂, 5 nM tritiated glutamic acid were added to the triplicate samples. The samples were incubated for thirty minutes at room temperature. The samples were harvested onto glass filters and were immediately washed with ice-cold 30 mM Tris, pH 7.0, 2.5 mM CaCl₂ under vacuum using an LKB 1295-001 automated cell harvester (Pharmacia LKB, Piscataway, NJ). The filters were dried in a microwave oven and counted in a gamma counter.

Protein determinations were carried out using a Coomassie Blue-based assay from Pierce Chemical Company (Rockford, IL) under conditions set forth by the manufacturer. One hundred microliters of undiluted cell homogenate or BSA standard was added to 2 ml of reagent and the optical density was measured at 595 nm. Protein concentrations of the samples were taken from a standard curve generated using the BSA standards diluted in 30 mM Tris, pH 7.0, 2.5 mM CaCl₂.

The results of these assays showed that quisqualate was able to competitively bind the glutamate receptor expressed by the transfected BHK cells.

Functional screening of agonists and antagonists

BHK 570 cells expressing GluGR or mock-transfected BHK 570 cells are plated into 24-well tissue culture dishes at about 100,000 cells per well. After 24 hours, the cells are labeled with 0.2 μCi of myo-(2-³H) inositol (specific activity - 20 Ci/mmol; New England Nuclear,) per w 11. At the nd of a 24 to 48 hour incubation, the cells are washed with prewarmed DMEM (Dulbecc 's Modified Eagles Medium; Product No. 51-432, JRH Biosci nces, Len xa, KS) which has been buffered t pH 7.4 with Hepes

buff r (Sigma Chemical Co.) containing 10 mM LiCl, and are incubat d for fiv minutes at 37. The selected drugs are then added and the cells are incubated for an additional thirty minutes at 37°C. The reaction is stopped by placing the cells on ice, and the cells are lysed by aspirating off the media and adding 0.5 ml of cold DMEM and 0.5 ml of ice-cold 10% perchloric acid. After ten minutes the cell lysate is transferred to a tube on ice containing 250 μ l 10 mM EDTA, pH 7.0. The samples are neutralized with 325 μ l of 1.5 M KOH in 60 mM Hepes Buffer. After the precipitates settles, 1.0 ml of the supernatant is applied to an Amprep minicolumn (Amersham, Arlington Heights, IL, RPN1908). Inositol phosphates are eluted off the column and samples are counted in a scintillation counter. A positive response is indicated by an increase in labeled inositol phosphate levels.

EXAMPLE II

20

25

30

35

15

5

10 .

Screening for additional glutamate receptor subtypes

Additional glutamate receptor subtypes were isolated using probes derived from clone 45-A. Glutamate receptor subtypes were isolated from a total rat brain cDNA library in Lambda Zap II, which was size selected for inserts of 3 kb before ligation (prepared for Terry Snutch, Ph.D., University of British Columbia, Vancouver, British Columbia, Canada by Stratagene Cloning Systems, La Jolla, CA) and a rat cerebellum cDNA library in Lambda Zap II, which was size selected for inserts of 3 kb before ligation (Stratagene Cloning Systems, La Jolla, CA).

The total rat brain library and the rat cerebellum library were plated out with <u>E. coli</u> XL-1 cells onto NZY agar plat s (Tabl 3) to obtain approximately 2.1 x 10⁶ plaques. Clone 45-A, encoding subtype la, was digested with Pst I to isolate the 1.3 and 1.6 kb fragments. The 45-A Pst I fragments wer labeled by random priming using

the Amersham random-priming kit (Am rsham, Arlington Hts, IL). Duplicate lifts were prepar d from the plates, and the filters were hybridized with the probes in 50% formamide at 37°C. After an overnight hybridization, the filters were washed in 2x SSC + 0.1% SDS at 50°C. Positive plaques were isolated by several rounds of dilution plating and repeated screening with the random-primed probes.

10

Table 3

NZY Agar

To 950 ml of deionized water, add:

10 g NZ amine: Casein hydrolysate enzymatic (ICN Biochemicals)

5 g NaCl

5 g bacto-yeast extract

1 g casamino acids

2 g MgSO₄ 7H₂O

20

15

Shake until the solutes have dissolved, Adjust to pH 7.0 with 5 N NaOH (approximately 0.2 ml). Adjust the volume of the solution to 1 liter with deionized $\rm H_2O$. Sterilize by autoclaving for 20 minutes.

25

20x SSC

Dissolve 175.3 g NaCl and 88.2 g sodium citrate in 800 ml $\rm H_2O$. Adjust the pH to 7.0 with a few drops of 10 N NaOH. Adjust the volume to 1 liter with $\rm H_2O$. Sterilize by autoclaving.

30

35

Plasmid DNA was prepared from positive plaques using the Bluescript system (Stratagene Cloning Systems). The plasmid DNA was subjected to restriction analysis and Southern blot analysis (Sambrook et al., ibid., which is incorporated herein by reference). Two clones, SN23, d rived from the total rat brain library, and SR2, derived from the rat cereb llum library, w re id ntified

15

20

25

30

35

as being different than the 45-A clone and were sequenced. Sequence analysis sh wed that they represented two additional subtypes. SN23 encodes subtype 1b, which contains an additional 85 bp exon that encodes a new stretch of 20 amino acids and a stop codon in the intracellular domain, is 292 amino acids shorter than the 45-A clone. The nucleotide sequence and deduced amino acid sequence of clone SN23 are shown in Fig. 7. SR2 was found to contain a partial cDNA sequence encoding subtype 2a, which is a novel sequence that shares a 42% homology to the transmembrane domains and extracellular domain of the 45-A clone.

A complete subtype 2a clone was obtained by rescreening both libraries as described above with the radiolabeled 1.3 kb Pst I fragment from clone 45-A and a radiolabeled 1.4 kb Eco RI-Pvu II fragment from SR2. Two additional clones were obtained. SN30, derived from the total rat brain library, contained the entire subtype 2a coding sequence. The nucleotide sequence and deduced amino acid sequence of clone SN30 are shown in Fig. 8. SR13, derived from the rat cerebellum library, contained an incomplete sequence of a new receptor subtype, 2b. Sequence analysis of SR13 showed that the coding sequence was incomplete at the 3' end and was virtually identical to the SN30 sequence except that it contained a 610 base pair deletion within the 3' terminus of SN30. The DNA sequence of the cDNA insert in clone SR13 is shown in Figure 9.

The complete 3' end of the subtype 2a clone was generated using PCR amplification and an oligonucleotide containing a sequence unique to SR13 (ZC4520, Table 4) and an oligonucleotide corresponding to a sequence near the 3' end of the 3' non-translated region of SN30 (ZC4519, Table 4). DNA was prepared from plate lysates of the original plating of ach library. Each plate pr duced a pool of clones. For the PCR reactions, ten nanograms from each library and 100 pmol of each oligonucleotid w r combin d in a reaction volume of 50

μl containing 50 mM KCl, 10 mM Tris-HCl pH 9.0, 1.5 mM MgCl₂, 0.1% Triton X-100, 0.01% gelatin, 0.2 mM each deoxynucleotide triphosphate and 2.5 units of Thermus aquaticus (Taq) DNA polymerase (Promega Corporation, Madison, WI). The reaction mixture was overlaid with mineral oil. After five cycles (30 seconds at 94°C, 30 seconds at 45°C, 1 minute at 50°C) and twenty-five cycles (30 seconds at 94°C, 30 seconds at 45°C, 1 minute at 72°C) the amplified DNA was removed for analysis.

10

15

20

25

Table 4

Degenerate Oligonucleotide Primer Sequences (5' - 3')

ZC4519

TTT ATT AGA AAT GTT CTC GGT

ZC4520

CCT CTT CCA TAT TTT TCC ATT

ZC4559

ATA AGA ATT CAT NKR YTT NGC YTC RTT RAA

ZC4560

ATA AGA ATT CTT YRA YGA RAA NGG NGA YGC

ZC4561

ATA AGA ATT CGC NGG NAT HTT YYT NKG NTA

ZC4562

ATA AGA ATT CTA NCM NAR RAA DAT NCC NGC

ZC4563

ATA AGA AAT CAN GTN GTR TAC ATN GTR AA

30

35

An aliquot from each reaction was electrophoresed on agarose and transferred to nitrocellulose for Southern analysis. Southern analysis of the PCR products showed that a 460 bp fragment corresponding to the 3' end of the 2b sequence was present in several pools. One of the pools that produced the correct size PCR product encoding the 3' sequence of the 2b subtype was diluted and screened with radiolabeled ZC4519 and ZC4520 (Table 4). Phage that hybridiz to both radiolab led ZC4519 and ZC4520 are picked, eluted, dilut d, plated and rescreened with the oligonucleotide probes. The screening is

10

15

20

25

30

35

repeated until a pure clon is obtain d. The pur clone is sequenced, and a full-length clon is constructed using the most convenient restriction enzyme(s).

Based on an alignment of the deduced amino acid sequences of subtypes 1a and 2a, strategies were designed for cloning additional subtypes using PCR amplification. Degenerate oligonucleotide families were prepared to encode conserved amino acid sequences in the sixth transmembrane domain, a region surrounding the conserved amino acid sequence Phe-Asp-Glu-Lys, the third cytoplasmic loop, and the second transmembrane domain (Table 4).

Glutamate receptor cDNA sequences were amplified with pairs of degenerate primers from Table 4 using the PCR method on cDNA from the total rat brain library, the cDNA from the rat cerebellum library, a rat cortex cDNA library or a rat hippocampus cDNA library (both obtained from Michael Brownstein, National Institutes of Health, Bethesda, MD). The primers also each contained a 5' tail of 10 nucleotides, which provided convenient restriction enzyme sites. For each PCR reaction, ten nanograms from the library and 100 pmol of the oligonucleotide pools ZC4563 and ZC4560 (Table 4) were combined in a reaction volume of 50 µl containing 50 mM KCl, 10 mM Tris-HCl pH 9.0, 1.5 mM MgCl2, 0.1% Triton X-100, 0.01% gelatin, 0.2 mM each deoxynucleotide triphosphate and 2.5 units of Taq DNA polymerase. The reaction mixture was overlaid with mineral oil. After five cycles (30 seconds at 94°C, 30 seconds at 45°C, 1 minute at 50°C) and twenty-five cycles (30 seconds at 94°C, 30 seconds at 45°C, 1 minute at 72°C) the amplified DNA was removed for analysis.

An aliquot from each reaction was electrophoresed on an agarose gel. Southern analysis of the gel was performed using essentially the method described by Sambrook et al. (ibid.) and random-primed fragments covering the entire coding r gions from both the subtype la and 2a clones. The autoradiographs showed that the PCR reaction gen rated fragments of novel size that wer

10

15

20

25

30

35

different from either the la or 2a subtype. PCR-generated fragments w re lectrophoresed on an agarose gel. Regions corresponding to the unique-sized receptor-related products were excised and electrophoresed onto NA45 paper (Schleicher and Schuell, Keene, NH). The purified fragments were recovered using essentially the method described by the manufacturer, digested with Eco RI and ligated to plasmid pVEGT' that had been linearized by digestion with Eco RI and treated with phosphatase to prevent recircularization. ligation mixtures were transformed into E. coli strain DH10b cells. Transformants were picked and replica plated onto nitrocellulose filters and screened using random-primed probes from the la and the la clones. Forty-eight colonies were picked for restriction analysis and sequencing.

DNA sequences from the cDNA from the total rat brain library and the cDNA from the rat cerebellum library were each amplified and analyzed using the methods described above and oligonucleotide ZC4559 in combination with either ZC4561 or ZC4559 (Table 4).

A rat cortex cDNA library and a rat hippocampus cDNA library (both obtained from Michael Brownstein, NIH) are subdivided into 30 pools of 10,000 colonies. DNA is prepared from each pool, and the DNA is subjected to Southern analysis after restriction digestion of the pools with Bam HI and Xho I or by PCR amplification of each pool using the degenerate oligonucleotides of Table The library pools containing DNA that hybridize to the probes and appear to contain a full-length cDNA are subdivided. The plasmid DNA is prepared and screened as described above. Positive pools are again divided and the procedure is continued until the pool is reduced to The clones are subjected to restriction pure clones. analysis and partial sequence analysis. Clones that represent distinct glutamate receptor homologs are compl t ly s qu nced. Full length clon s ar g n rated by subjecting the original pools to PCR amplification

using an oligonucleotide primer specific to the SP6 promoter at the 5' end of the cDNA insert and an antisense oligonucleotide primer corresponding to the 5' end of the most complete cDNA to identify pools that contain the longest glutamate receptor homolog cDNA. The pool is then diluted and rehybridized with the probes as described above to isolate a full length cDNA clone.

Expression of Glutamate Receptor Subtypes

10

15

20

5

Complementary DNA sequences encoding subtypes 1b and 2a were subcloned first into the mammalian expression vector Zem228R to obtain convenient terminal restriction The cDNAs were then subcloned into pVEGT'. cDNA sequence encoding subtype 1b was constructed by replacing the 3' terminal portion of subtype la described in Example I with the analogous portion of subtype 1b from SN23. Plasmid SN23 was digested with Kpn I and Xba I to isolate the fragment containing the 3' terminus of the 1b subtype. The plasmid containing the subtype 1a coding sequence (45-A) in Zem228R was digested with Kpn I and Xba I to isolate the vector containing fragment. The vector containing fragment is ligated to the Kpn I-Xba I fragment from SN23. The resulting plasmid comprises the MT-1 promoter, the subtype 1b cDNA and the hGH terminator. This plasmid was transfected into the BHK 570 cell line essentially as described in Example I to obtain stably transfected cell lines expressing the subtype 1b receptor. The subtype 1b cDNA fragment was isolated as a Bam HI fragment, which was ligated with pVEGT' that had been linearized with Bam HI. A plasmid containing the cDNA sequence in the correct orientation was used to synthesize RNA in an in vitro system. The RNA was injected into occytes as described above.

30

35

25

Plasmid SN30, which comprises the subtype 2a cDNA, was digested with Eco RI to isolate the subtype 2a cDNA. Th Eco RI fragment was ligated with Eco RI-lin ariz d Zem228R. A plasmid containing the insert in th correct ori ntation was digest d with Bam HI to isolate the cDNA

10

15

20

sequence. The Bam HI fragment comprising the subtype 2a cDNA was light d with Eco RI-lineariz d pVEGT'. A plasmid containing the cDNA in the correct orientation was used to synthesize RNA in an in vitro translation. The RNA was injected into frog oocytes as described above.

EXAMPLE III

Generation of antibodies to glutamate receptor subtypes

Receptor subtype-specific polyclonal antisera were generated in rabbits using standard immunization techniques. Synthetic peptides (Table 5) were designed from the cloned receptor sequences. The peptides were conjugated to keyhole limpet hemocyanin, and each antigen was used to immunize two animals. For each peptide, the animals were injected with $100-200~\mu g$ of conjugated peptide divided among three subcutaneous sites. The animals were immunized at three-week intervals and bled via an ear vein 10 days after the third and subsequent immunizations.

Table 5

Subtype 25	Seq. ID	Peptide <u>Sequence</u>	Apparent Location
25 1a	21	RDSLISIRDEKDGLNRC	extracellular
	22	DRLLRKLRERLPKARV	extracellular
*	23	EEVWFDEKGDAPGRYD	extracellular
	24	EFVYEREGNTEEDEL	cytoplasmic
30	25	PERKCCEIREQYGIQRV	extracellular
	26	IGPGSSSVAIQVQNLL	extracellular
	27	IAYSATSIDLSDKTL	extracellular
1b	28	KKPGAGNAKKRQPEFS	cytoplasmic
	29	PEFSPSSQCPSAHAQL	cytoplasmic
35 2a	30	DKIIKRLLETSNARG	extracellular
	31	VNFSGIAGNPVTFNEN	extracellular
	32	GEAKSELCENLETPAL	cytoplasmic
2b	33	PARLALPANDTEFSAWV	cytoplasmic

Anti-peptide antibodies w r purified by affinity purification using the ProtonTM Kit (Multiple Peptide Systems (San Diego, CA). Purified antibodies were stored in column elution buffer and neutralizing buffer (supplied by Multiple Peptide Systems). Bovine serum albumin was added to a concentration of 1 mg/ml, and sodium azide was added to a concentration of 0.05%. The antibodies were stored at 4°C or in small aliquots at -20°C.

10

Antibodies generated from the peptides listed in Table 6 were used to detect G protein-coupled glutamate receptors by Western blot analysis of membranes prepared from transfected cell lines that were stably expressing the subtype 1a or subtype 1b receptors. Control cell lines were transfected with vector alone.

15

Table 6

Analysis of Antibodies Raised to Peptides

20	Antibodies to Peptide Sequence	Seq. ID	Location	Western
	RDSLISIRDEKDGLNRC	21	extracellular	+++ with bkgd
	DRLLRKLRERLPKARV	22	extracellular	+
25	EEVWFDEKGDAPGRYD	23	extracellular	++++ low bkgd
	EFVYEREGNTEEDEL	24	cytoplasmic	++++ low bkgd
Ð	KKPGAGNAKKRQPEFS	28	cytoplasmic	+ for la
× .	*		· · · · · · · · · · · · · · · · · · ·	- for 1b
	PEFSPSSQCPSAHAQL	29	cytoplasmic	+++ for 1b low bkgd

30

35

40

Transfectants that were stably expressing either the 1a or 1b subtype were each grown to confluency in five to ten 150 mm plates. Each plate was first washed twice with 15 ml of cold PBS and then 20 ml of ice cold 10 mM NaHCO3 was added to each plate. The cells from each plate were scraped off the plates with a rubb r spatula and transferr d to a glass dounc h mogeniz r on ice. The cells w re disrupted with ten strokes of the B pestle. Th homog nates from ach plate were combined

10

15

20

25

30

35

and centrifuged for thirty minutes at 3000 rpm at 4°C. The pellets w re resuspend d in 4-8 ml of 10 mM NaHCO, using a 22 g needle and syringe, and 69% sucrose was added (6-12 ml) to each sample until an index of refraction of 1.410 was reached. The samples were transferred to a high speed centrifugation tube, and each sample was overlayed with 42% sucrose. The samples were centrifuged for two hours at 25,000 rpm at 4°C. The samples were collected by gently floating the membranes off the 42% sucrose layer by adding 1 ml of 10 mM NaHCO, and resuspending the membranes by carefully stirring the upper layer. The upper layer was transferred to a fresh tube on ice. The purified membranes were centrifuged at 10,000 rpm at 4°C and the pellets resuspended in 10 mM NaHCO,. The purified membranes were then adjusted to a final protein concentration of 1-2 μ g/ml.

Ten to twenty micrograms of each purified membrane preparations were diluted with 2x SDS-mercaptoethanol buffer (100 mM Tris HCl (pH 6.8), 200 mM dithiothreitol, 4% SDS, 0.2% bromophenol blue, 20% glycerol). The samples were incubated for 15 minutes at 37°C followed by boiling for 5 minutes. The samples were subjected to SDS-PAGE on 4-15% gradient gel. The samples were electrotransferred to nitrocellulose using the method essentially described by Towbin (Proc. Natl. Acad. Sci. USA 76: 4350-4354, 1979; which is incorporated herein by reference in its entirety). After transfer, the nitrocellulose was cut into strips such that each strip contained a control and receptor samples. nitrocellulose was preincubated in blocking buffer and then incubated with a dilution of either the preimmune serum or the serum collected after antigenic stimulation (serum from later bleeds (i.e. those after four antigen stimulations) were diluted 1:1500). After washing, a horse radish peroxidase-conjugated goat anti-rabbit antibody (Bio-Rad Laboratories, Richmond, CA) dilut d 1:2,500 was added and after incubation and washing, the horse radish peroxidase substrate (Bio-Rad Laboratories)

WQ 92/10583 PCT/US91/09422

66

was add d and the color reaction was initiated. The reaction was stopped by rinsing the filters in distilled water. Table 6 shows the results of the Western blot analysis.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Mulvihill, Eileen R. Hagen, Frederick S. Houamed, Khaled M. Almers, Wolfhard
- (ii) TITLE OF INVENTION: G PROTEIN-COUPLED GLUTAMATE RECEPTORS
- (iii) NUMBER OF SEQUENCES: 33
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Townsend and Townsend
 - (B) STREET: One Market Plaza, Steuart Street Tower
 - (C) CITY: San Francisco
 - (D) STATE: California
 - (E) COUNTRY: USA
 - (F) ZIP: 94105-1492
 - (V) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (Vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 07/672,007
 - (B) FILING DATE: 18-MAR-1991
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 07/648,481
 - (B) FILING DATE: 30-JAN-1991
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 07/626,806
 - (B) FILING DATE: 12-DEC-1990
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Parmelee, Steven W.
 - (B) REGISTRATION NUMBER: 31,990
 (C) REFERENCE/DOCKET NUMBER: 13952-6PC
 - (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (206) 467-9600
 - (B) TELEFAX: (415) 543-5043
- (2) INFORMATION FOR SEQ ID NO:1:

. 697

(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 4300 base pairs	
(A) LENGTH: 4300 Dabb p	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: SINGLE	
(D) TOPOLOGY: linear	• 111
(B) 102020C1	
	•
(ii) MOLECULE TYPE: CDNA	
COURCE!	•
(Vii) IMMEDIATE SOURCE:	
(B) CLONE: 45-A	i i
ALLA DESMITER	. :
(ix) FEATURE:	
(A) NAME/KEY: CDS	
(B) LOCATION: 3773973	•
	8
(vi) SEQUENCE DESCRIPTION: SEQ ID NO:1:	
(vi) SEQUENCE DESCRIPTION. DEC	
CCGAGAACGG CTGCAGTCCT CTGACCTGAG ACCAATAGCT GTGTCTACCC GGACTCAGCG	60
CCACAACGG CTGCAGTCCT CTGACCTGAG ACCAATAGCT GIGIGIA	
CCGAGAACGG CIGGIGGCCCCCA	120
TCCAGCTCAC CGCCACTAAC GCGCCGCGCA TTGGACACCT GATCCACACA CCTTCGGGCA	
TCCAGCTCAC CGCCACTAAC GCGCCGCGCA III	
CCAGTGAAAA ACCGCGACTT GATTTTCTGG AAGAACGCCC CCAGGGTGTG GGAGCGGTCG	180
COLORGIANA ACCOCGACTT GATTTTCTGG AAGAACGCCC CCAGGGTGT	* -
CCAGTGAAAA ACCGCCT	240
CCCACCGA GAGGGGCAGT AGTGGAGGCA GAGAAAGCGI	240
TGGAGGACCA GCAGGAGGAA GCGGAGGGGA GAGGGGGCAGT AGTGGAGGCA GAGAAAGCGT	
TGAACCAGCT GTGTTGGCCG AAGGCACGAA ACGGCAAAAG GCAGCGGTGA GCATCTGTGT	300
TO LOCA COT CTCTTGGCCG AAGGCACGAA ACGGCAAAAG GCAGCGCCC	
TGAACCAGCI GIGIIGGGCCCCCA	360
GGTTCCCGCT GGGAACCTGC AGGCAGGACC GGCGTGGGAA CGTGGCTGGC CCGCGGTGGA	. 300
GGTTCCCGCT GGGAACCTGC AGGCAGGACC CLOUD	
CCGCGTCTTC GCCACA ATG GTC CGG CTC CTC TTG ATT TTC TTC CCA ATG	409
COCCONOMIC COCACA ATG GTC CGG CTC CTC TIG ALL TIG BEC BEC Met	
CCGCGTCTTC GCCACA ATG GTC CGG CTC CTC 11G ATT Met Val Arg Leu Leu Leu Ile Phe Pro Met 10	- 1
Met var 5	٠,
ATC TTT TTG GAG ATG TCC ATT TTG CCC AGG ATG CCT GAC AGA AAA GTA	457
THE MAN MAN CAG ATG TCC ATT TTG CCC AGG ATG CCT AND AND THE VAL	
ATC TTT TTG GAG ATG TCC ATT TTG CCC AGG ATG CCT AND ANG Lys Val Ile Phe Leu Glu Met Ser Ile Leu Pro Arg Met Pro Asp Arg Lys Val 20	-00
Ile Phe Leu Glu Met Sel 120 20	
13	
TTG CTG GCA GGT GCC TCG TCC CAG CGC TCC GTG GCG AGA ATG GAC GGA	505
TO GO COT GCC TCG TCC CAG CGC TCC GTG GCG AGA NOT AGA GIV	
TTG CTG GCA GGT GCC TCG TCC CAG CGC TCC GTG GCG ASS MET ASP GLY Leu Leu Ala Gly Ala Ser Ser Gln Arg Ser Val Ala Arg Met Asp Gly 35	
Leu Leu Ala Gly. Ala Sel St. 35	•
30	
THE CAR COT CCA GCC	553
GAT GTC ATC ATC GGA GCC CTC TTC TCA GTC CAT CAC CAG CCT CCA GCC	.1
GAT GTC ATC ALL SAN Phe Ser Val His His GIN Pro PIO ALL	
Asp Val Ile Ile Gly Ale 200 Inc.	_
A5	
THE CALL THE GET	601
GAG AAG GTA CCC GAA AGG AAG TGT GGG GAG ATC AGG GAA CAG TAT GGT	•
GAG AAG GTA CCC GAA AGG AAG TGT GGG GAG ATC AGG GAU GIN TYR GIY Glu Lys Val Pro Glu Arg Lys Cys Gly Glu Ile Arg Glu Gln Tyr Gly 75	
Glu Lys Val Pro Glu Rig 210 11 70	
an Kina na na arawan na katao na wakata na katao na na katao na katao na manaka na katao na manaka na katao na	
THE NAME OF STATE AND SACE OF STATE AND SACE OF SACE O	649
ATC CAG AGG GTG GAG GCC ATG TTC CAC ACG TTG GAT AAG ATT AAC GCG	
ATC CAG AGG GTG GAG GCC ATG TTC CAC ACG ITG GAT Lys Ile Asn Ala Ile Gln Arg Val Glu Ala Met Phe His Thr Leu Asp Lys Ile Asn Ala 90	*
Tie Gin Arg Val Giu Ara Mos III	
80	

GAC CCG GTG CTC CTG CCC AAC ATC ACT CTG GGC AGT GAG ATC CGG GAC ASP Pro Val Leu Pro Asn Ile Thr Leu Gly Ser Glu Ile Arg Asp

TCC TGC TGG CAC TCT TCA GTG GCT CTC GAA CAG AGC ATC GAA TTC ATC Ser Cys Trp His Ser Ser Val Ala Leu Glu Gln Ser Ile Glu Phe Ile 110 AGA GAC TCC CTG ATT TCC ATC CGA GAT GAG AAG GAT GGG CTG AAC CGA ATG ASP Ser Leu Ile Ser Ile Arg Asp Glu Lys Asp Gly Leu Asn Arg 125 TGC CTG CCT GAT GGC CAG ACC CTG CCC CCT GGC AGG ACT AAG AAG CCT Cys Leu Pro Asp Gly Gln Thr Leu Pro Pro Gly Arg Thr Lys Lys Pro 140 ATT GCT GGA GTG ATC GGC CCT GGC TCC AGC AGG ACT AAA AAG CCT Cys Leu Pro Asp Gly Gln Thr Leu Pro Pro Gly Arg Thr Lys Lys Pro 140 ATT GCT GGA GTG ATC GGC CCT GGC TCC AGC AGG ACT AAA AAG CCT Cys Leu Pro Asp Gly Fro Gly Ser Ser Ser Val Ala Ile Gln Val 160 ATT GCT GGA GTG ATC GGC CTT GGC AGC TCT GTG GCC ATT CAA GTC Ile Ala Gly Val Ile Gly Pro Gly Ser Ser Ser Val Ala Ile Gln Val 160 CAG AAT CTT CTC CAG CTG TTC GAC ACT CCA CAG ATC GCC ATT TCT GCC Gln Asn Leu Leu Gln Leu Phe Asp Ile Pro Gln Ile Ala Tyr Ser Ala 185 ACA AGC ATA GAC CTG AGT GAC AAA ACT TTG TAC AAA TAC TTC CTG AGG Thr Ser Ile Asp Leu Ser Asp Lys Thr Leu Tyr Lys Tyr Phe Leu Arg 190 GTG GTC CCT TCT GAC ACT TTG CAG GCA AGG GCA ATG CTC GAC ATA GTC Val Val Pro Ser Asp Thr Leu Gln Ala Arg Ala Het Leu Asp Ile Val 205 AAG CCT TAC AAC TGG ACC TAT GTC TCA GAC GTC CAC ACA GAA GGC AAT 1081 Lys Arg Tyr Asn Trp Thr Trr Val Ser Ala Val His Thr Glu Gly Asn 1081 Lys Arg Tyr Asn Trp Thr Trr Val Ser Ala Val His Thr Glu Gly Asn 1081 CGC CTC TGC AGC AGA TGG ATG GAT TCT AAC CAC ACA GAA GGA ATT 1081 GGC CTC TGC AGC CAC CAC TCG GAC AAA CTC AGC AGC ACT GGC GAC 1177 GCC AGG GTT GTG CCC CAC CTC CTC CTC AAA CTC CCC AGC CTC CAC AGA AGC CTT GAC ACT GGA ATG GTC TCC AAC ACA ACA ACA GAC GAC 1225 CCC AGG GTT GTG GTC TCC TCC CTC CCC AAA CTC GCC GAC CTC CAC ALA ATG GLy Ser Phe Asp Arg Leu Leu Arg Lys Ile Tyr Ser Asn Ala Gly Glu 275 CCC AGG GTT GTG GTC TCC TCC CTC GCC GAC GCC ATG CAC CTC CAC ALA ATG GLY Ser Phe Asp Arg Leu Leu Arg Lys Leu Arg Glu Arg Leu Pro Lys 1295 CCG AGG GTT GTG GTC TCC TCC CTC GCC GTC GCC GC			,,,,	•	95			• •		100				•	105	• •		•	
See Cys Trp His Ser Ser Val 115 120					•		0	•									A THIC		745
135 136 136 135		TCC Ser	TGC Cys	Trp	CAC	TCT Ser	TCA (Ser \	AGT 4	WT a	Leu CTC	GAA Glu	CAG Gln			GAA Glu	Phe	Ile		*
TGC CTG CCT GAT GGC CAG ACC CTG CCC CCT GGC AGG ACT AAG AAG CCT CYS Leu Pro Asp Gly Gln Thr Leu Pro Pro Gly Arg Thr Lys Lys Pro 140 ATT GCT GGA GTG ATC GGC CCT GGC TCC AGC TCT GTG GCC ATT CAA GTC 150 ATT GCT GGA GTG ATC GGC CCT GGC TCC AGC TCT GTG GCC ATT CAA GTC 160 ATT GCT GGA GTG ATC GGC CTG GTC GAC ATC CCA AGA ATC GCC ATT CAA GTC CAC AAT CTT CTC CAG CTG TTC GAC ATC CCA CAG ATC GCC TT TCT GCC GTG Nasn Leu Leu Gln Leu Phe Asp 11e Pro Gln 11e Ala Tyr Ser Ala 185 ACA AGC ATA GAC CTG AGT GAC AAA ACT TTG TAC AAA TAC TTC CTG AGG 195 Thr Ser 11e Asp Leu Ser Asp Lys Thr Leu Tyr Lys Tyr Phe Leu Arg 190 GTG GTC CCT TCT GAC ACT TTG CAG GCA AGG GCG ATC CTC GAC ATA GTC 1033 Val Val Pro Ser Asp Thr Leu Gln Ala Arg Ala Met Leu Asp 11e Val 205 AAG CGT TAC AAC TGG ACC TAT GTC TCA GCA GTC CAC ACA GAA GGG AAT 1081 Lys Arg Tyr Asn Trp Thr Tyr Val Ser Ala Val His Thr Glu Gly Asn 235 TAC GGC GAG AGT GGA ATG GAT GCT TTC AAA GAC CTG GCC CAG GAA T 1081 Tyr Gly Glu Ser Gly Met Asp Ala Phe Lys Glu Leu Ala Ala Glu Glu 250 GCC CTC TGC ATC GCA CAC TCG GAC AAA ATC TAC AGC AAT GCT GCC CAG GAC AGA GGY GIV GLU CYS 11e Ala His Ser Asp Lys 11e Tyr Ser Asn Ala Gly Glu 255 AAG AGC TTT GAC CGC CTC CTC CTC CTC AAA ATC TCC GG GAC CTC CCC AAG 1177 GCC AGG GTT GTG GTC TCT CTC TCC GAG GGC ATG CTC GCC CAG GAC 1177 GCC AGG GTT TGT GTC TCC TTC TCC GAG GGC ATG CTC GGG GGC TTA ALA Arg Lau Val Val Val Cys Phe Cys Glu Gly Het Thr Val Arg Gly Leu 2290 GCC AGG GTT GTG GTC TCC TTC TCC GAG GGC ATG ACA GTG CGG GGC TTA ALA Arg Val Val Val Val Cys Phe Cys Glu Gly Het Thr Val Arg Gly Leu 2290 GCC AGG GTT GTG GTC TCC TTC TCC GAG GGC ATG ACA GTG CGG GGC TTA ALA Arg Val Val Val Cys Phe Cys Glu Gly Het Thr Val Arg Gly Leu 2290 GCC AGG GTT GTG GTC TCC TTC TCC GAG GGC ATG ACA GTC CGG GGC TTA ALA Arg Leu Gly Val Val Val Gly Glu Ph Ser Leu Ile 100 GCA AGT GAT GGA TGG GCC GCC CTG GGC GTC GTC GCC GAG TTC CCA ATT 1321 GCA AGT GAT GGA TGG GCC GCC GTC GTC GTC GTC GCC GAG TTC TAT GAG GLy Ser Asp Gly Trp Ala Asp Arg Asp Glu Val Ile Glu Gly Tyr Glu 315		AGA Arg	Asp	TCC	CTG Leu	ATT Ile	Ser .	TTE	CGA Arg	GAT Asp	GAG Glu	AAG Lys	GAT Asp 135	GGG Gly	CTG Leu	AAC Asn	CGA Arg	, ,	793
ATT GCT GGA GTG ATC GGC CCT GGC TCC AGC TCT GTG GCC ATT CAA GTC Ile Ala Gly Val Ile Gly Pro Gly Ser Ser Ser Val Ala Ile Gln Val 160 CAG AAT CTT CTC CAG CTG TTC GAC ATC CCA CAG ATC GCC TAT TCT GCC GIn Asn Leu Leu Gln Leu Phe Asp Ile Pro Gln Ile Ala Tyr Ser Ala 185 ACA AGC ATA GAC CTG AGT GAC AAA ACT TTG TAC AAA TAC TTC CTG AGG PTC TCT GCC AST Leu Asp Leu Ser Asp Lys Thr Leu Tyr Lys Tyr Phe Leu Arg 195 GTG GTC CCT TCT GAC ACT TTG CAG GCA AGG GCG ATG CTC GAC ATA GTC Val Val Pro Ser Asp Thr Leu Gln Ala Arg Ala Met Leu Asp Ile Val 205 AAG CGT TAC AAC TGG ACC TAT GTC TCA GCA GTC CAC ACA GAA GGC AAT 1081 Lys Arg Tyr Asn Trp Thr Tyr Val Ser Ala Val His Thr Glu Gly Asn 230 GCG GTG GGC GAG AGT GGA ATG GAT GCT TTC AAA GAA CTG GCT GCC CAG GAA 1129 TAC GGC GAA AGT GGA ATG GAT GCT TTC AAA GAA CTG GCT GCC CAG GAA 1129 GCC CTC TGC ATC GCA CAC TCG GAC AAA ACC TAC AGC AAT GTC GCT GCC CAG GAC ACA GAA AAA ACC TAC GAA ACC AACA GAA AAA ACC TAC GCA GAC GAA AAA ACC TAC GCA GAC GAC GAC GAC GAC GAC GAC GAC G		Cys		CCT Pro	GAT Asp	GJÄ'	CAG		CTG Leu	CCC Pro	CCT Pro	GIY	AGG Arg	ACT Thr	AAG Lys	AAG Lys	CCT Pro 155		841
CAG AAT CTT CTC CAG CTG TTC GAC ATC CCA CAG ATC GCC TAT TCT GCC Gln Asn Leu Leu Gln Leu Phe Asp 11e Pro Gln 11e Ala Tyr Ser Ala 175 ACA AGC ATA GAC CTG AGT GAC AAA ACT TTG TAC AAA TAC TTC CTG AGG Thr Ser 11e Asp Leu Ser Asp Lys Thr Leu Tyr Lys Tyr Phe Leu Arg 190 GTG GTC CCT TCT GAC ACT TTG CAG GCA AGG GCG ATG CTC GAC ATA GTC 190 CTG GTC CCT TCT GAC ACT TTG CAG GCA AGG GCG ATG CTC GAC ATA GTC 200 AAG CGT TAC AAC TGG ACC TAT GTC TCA GCA GTC CAC ACA GAA GGG AAT Lys Arg Tyr Asn Trp Thr Tyr Val Ser Ala Val His Thr Glu Gly Asn 220 TAC GGC GAG AGT GGA ATG GAT GCT TTC AAA GAA CTG GCT GCC CAG TYr Gly Glu Ser Gly Met Asp Ala Phe Lys Glu Leu Ala Ala Gln Glu 255 AAG AGC TTT GAC CGC CAC CTG GAC AAA ATC TAC AGC AAT GCT GGC GAG GLY Leu Cys 11e Ala His Ser Asp Lys 11e Tyr Ser Asn Ala Gly Glu 255 AAG AGC TTT GAC CGG CTC CTG CGT AAA CTC CGG GAG CGT CTC CCC AAG 1127 GCC AGG GTT GTG GTC TCC TTC TGC GAG GGC ATG ACA GTG CGG GGC TTA ALA ACT TGT GAC CGG CTC CTG CGT AAA CTC CGG GAG CGT CTC CCC AAG 1225 CCC AGG GTT GTG GTC TCC TTC TGC GAG GGC ATG ACA GTG CGG GGC TTA ALa Act Val Val Val Cys Phe Cys Glu Gly Met Thr Val Act Gly Leu 285 CTG AGT GCC ATG CGC CGC CTG GGC GTG GGC GAG GTT CTC ATT 1321 CGG AGT GAT GGA TGG GCA GAC AGA GAT GAA GTC TTC CAC ATT 1321 CTG AGT GCC ATG GGC CTG GGC GTG GGC GTG GGC TAT GAG ALS ACT GGT GGA GGA AGA GAT GAA GTC TTC CTC ATT 1321 CTG AGT GCC ATG GGC CTG CTG GGC GTG GGC GAG TTC TCA CTC ATT 1321 CTG AGT GCC ATG GGC CTG GGC GTG GGC GTG TTC TCA CTC ATT 1321 CTG AGT GGA TGG GCA GAC AGA GAT GAA GTC ATC GAA GGC TAT GAG GLY Ser Abp Gly TTP Ala Asp Act Gly Val Val Gly Glu Ph Ser Leu Ile 300 GGA AGT GAT GGA TGG GCA GAC AGA GAT GAA GTC ATC GAA GGC TAT GAG GLY Ser Abp Gly TTP Ala Asp Act Gly Val Val Ile Gly Gly Tyr Glu 3330			GCT Ala	GGA Gly	GTG Val	Ile		CCT Pro	GGC Gly	2 et	Set	TCT Ser	GTG Val	GCC Ala	ATT	CAA Gln 170	GTC Val	* · •	889
Thr Ser Ile Asp Leu Ser Asp 195 GTG GTC CCT TCT GAC ACT TTG CAG GCA AGG GCG ATG CTC GAC ATA GTC 1033 Val Val Pro Ser Asp Thr Leu Gln Ala Arg Ala Met Leu Asp Ile Val 205 AAG CGT TAC AAC TGG ACC TAT GTC TCA GCA GTC CAC ACA GAA GGG AAT 1081 Lys Arg Tyr Asn Trp Thr Tyr Val Ser Ala Val His Thr Glu Gly Asn 235 TAC GGC GAG AGT GGA ATG GAT GGT TTC AAA GAA CTG GCT GCC CAG GAA 1129 TAC GGC CTC TGC ATC GCA CAC TCG GAC AAA ATC TAC AGC AAT GGT GGC GAG GIV Leu Cys Ile Ala His Ser Asp Lys Ile Tyr Ser Asn Ala Gly Glu 255 AAG AGC TTT GAC CGG CTC CTG CGT AAA CTC CGG GAG CGG CTT CCC AAG 1225 AAG AGC TTT GAC CGG CTC CTG CGT AAA CTC CGG GAG CGG CTT CCC AAG 1225 GCC AGG GTT GTG GTC TGC TTC TGC GAG GGC ATG ACA GTC CGG GGC TTA Ala Arg Val Val Val Cys Phe Cys Glu Gly Het Thr Val Arg Gly Leu 295 CTG AGT GCC ATG CGC CGC CTG GGC GTC GTG GGC GAG TTC TCA ATT 1321 Leu Ser Ala Met Arg Arg Leu Gly Val Val Gly Glu Ph Ser Leu Ile 300 GGA AGT GAT GGA TGG GCA GAC AGA GAT GAA GTC ATC GAA GGC TAT GAG GLy Ser Asp Gly Trp Ala Asp Arg Asp Glu Val Ile Glu Gly Tyr Glu 310	•••	CAG Gln	AAT Asn	CTT	Leu		CTG Leu	TTC Phe	GAC Asp	TTA	CCA Pro	CAG Gln	ATC Ile	GCC Ala	-3-	TCT Ser	GCC Ala		937
CTC CCT TCT GAC ACT TTG CAG GCA AGG GCG ATG CTC GAC ATA GTC CTC CTC CTC CAS ASP Thr Lau Gln Ala Arg Ala Met Lau Asp Ila Val Val Val Pro Ser Asp Thr Lau Gln Ala Arg Ala Met Lau Asp Ila Val Val Val Ser Ala Met Lau Asp Ila Val Asp Ila Val Asp Ila Val Val Asp Ila Thr Glu Glu	٠.	ACA Thr	AGC Ser	Ile	Asp	CTG Leu	AGT Ser	GAC Asp	AAA Lys 195	ACT	TTG Leu	TAC Tyr	AAA Lys		TTC Phe	CTG Leu	AGG Arg	0	985
AAG CGT TAC AAC TGG ACC TAT GTC TCA GCA GTC CAC ACA GAA GGG AAT Lys Arg Tyr Asn Trp Thr Tyr Val Ser Ala Val His Thr Glu Gly Asn 235 TAC GGC GAG AGT GGA ATG GAT GGT TTC AAA GAA CTG GCT GCC CAG GAA Tyr Gly Glu Ser Gly Met Asp Ala Phe Lys Glu Leu Ala Ala Gln Glu 240 GGC CTC TGC ATC GCA CAC TCG GAC AAA ATC TAC AGC AAT GCT GGC GAG Gly Leu Cys Ile Ala His Ser Asp Lys Ile Tyr Ser Asn Ala Gly Glu 255 AAG AGC TTT GAC CGG CTC CTG CGT AAA CTC CGG GAG CGG CTT CCC AAG Lys Ser Phe Asp Arg Leu Leu Arg Lys Leu Arg Glu Arg Leu Pro Lys 275 GCC AGG GTT GTG GTC TGC TTC TGC GAG GGC ATG ACA GTG CGG GGC TTA Ala Arg Val Val Val Cys Phe Cys Glu Gly Het Thr Val Arg Gly Leu 285 CTG AGT GCC ATG CGC CGC CTG GGC GTC GTG GGC GAG TTC TCA CTC ATT Lau Ser Ala Met Arg Arg Leu Gly Val Val Gly Glu Ph Ser Leu Ile 300 GGA AGT GAT GGA TGG GCA GAC AGA GAT GAA GTC ATC GAA GGC TAT GAG GIy Ser Asp Gly Trp Ala Asp Arg Asp Glu Val Ile Glu Gly Tyr Glu 330		GTG Val	Val	CCT		GAC Asp	ACT Thr	Leu	CAG Gln	GCA Ala	AGG Arg	GCG Ala		CTC	GAC Asp	ATA	GTC Val	ķ:-	1033
TAC GGC GAG AGT GGA ATG GAT GGT GCT TTC AAA GAA CTG GCT GCC CAG GAA TYF Gly Glu Ser Gly Met Asp Ala Phe Lys Glu Leu Ala Ala Gln Glu 245 GGC CTC TGC ATC GCA CAC TCG GAC AAA ATC TAC AGC AAT GCT GGC GAG Gly Leu Cys Ile Ala His Ser Asp Lys Ile Tyr Ser Asn Ala Gly Glu 255 AAG AGC TTT GAC CGG CTC CTG CGT AAA CTC CGG GAG CGG CTT CCC AAG Lys Ser Phe Asp Arg Leu Leu Arg Lys Leu Arg Glu Arg Leu Pro Lys 275 GCC AGG GTT GTG GTC TGC TTC TGC GAG GGC ATG ACA GTG CGG GGC TTA Ala Arg Val Val Val Cys Phe Cys Glu Gly Het Thr Val Arg Gly Leu 290 CTG AGT GCC ATG CGC CGC CTG GGC GTC GTG GGC GAG TTC TCA CTC ATT Leu Ser Ala Met Arg Arg Leu Gly Val Val Gly Glu Ph Ser Leu Ile 300 GGA AGT GAT GGA TGG GCA GAC AGA GAT GAA GTC ATC GAA GGC TAT GAG Gly Ser Asp Gly Trp Ala Asp Arg Asp Glu Val Ile Glu Gly Tyr Glu 330		Lys	CGT		AAC Asn	TGG	THE	TAT Tyr	GTC Val	TCA Ser	GCA Ala			ACA Thr	GAA Glu	GGG Gly	AAT Asn 235		1081
AAG AGC TTT GAC CGG CTC CTG CGT AAA CTC CGG GAG CGG CTT CCC AAG Lys Ser Phe Asp Arg Leu Leu Arg Lys Leu Arg Glu Arg Leu Pro Lys 270 GCC AGG GTT GTG GTC TGC TTC TGC GAG GGC ATG ACA GTG CGG GGC TTA Ala Arg Val Val Val Cys Phe Cys Glu Gly Het Thr Val Arg Gly Leu 280 CTG AGT GCC ATG CGC CGC CTG GGC GTC GTG GGC GAG TTC TCA CTC ATT Leu Ser Ala Met Arg Arg Leu Gly Val Val Gly Glu Ph Ser Leu Ile 300 GGA AGT GAT GGA TGG GCA GAC AGA GAT GAA GTC ATC GAA GGC TAT GAG Gly Ser Asp Gly Trp Ala Asp Arg Asp Glu Val Ile Glu Gly Tyr Glu 330				GAG	AGT Ser	GIA	Met	GAT Asp	GCT Ala	TTC Phe	. Lys	924	CTG Leu	GCT	GCC Ala	Glr 250	GAA Glu		1129
AAG AGC TTT GAC CGG CTC CTG CGT AAA CTC CGG GAG CGG CTT CCC AAG Lys Ser Phe Asp Arg Leu Leu Arg Lys Leu Arg Glu Arg Leu Pro Lys 270 GCC AGG GTT GTG GTC TGC TTC TGC GAG GGC ATG ACA GTG CGG GGC TTA Ala Arg Val Val Val Cys Phe Cys Glu Gly Het Thr Val Arg Gly Leu 290 CTG AGT GCC ATG CGC CGC CTG GGC GTC GTG GGC GAG TTC TCA CTC ATT Leu Ser Ala Met Arg Arg Leu Gly Val Val Gly Glu Ph Ser Leu Ile 300 GGA AGT GAT GGA TGG GCA GAC AGA GAT GAA GTC ATC GAA GGC TAT GAG Gly Ser Asp Gly Trp Ala Asp Arg Asp Glu Val Ile Glu Gly Tyr Glu 330		GGC	CTC	TGC Cys	; Ile	GCA Ala		TCG Ser	GAC Asp	, rås	TIPE	TAC	AGC Ser	AAT Asn	GCI Ala 265	GGG	GAG Glu		1177
Ala Arg Val Val Val Cys Phe Cys Glu Gly 1295 285 CTG AGT GCC ATG CGC CGC CTG GGC GTC GTG GGC GAG TTC TCA CTC ATT Leu Ser Ala Met Arg Arg Leu Gly Val Val Gly Glu Ph Ser Leu Ile 315 300 GGA AGT GAT GGA TGG GCA GAC AGA GAT GAA GTC ATC GAA GGC TAT GAG GGY Ser Asp Gly Trp Ala Asp Arg Asp Glu Val Ile Glu Gly Tyr Glu 330		AAG Lys	s Sei	: Pho	r GAC e Asi		CTC Leu	CTG Leu	Win	و برد	CTC Lev	CGG Arg	GAG Glu		•	CCC Pro	C AAG D Lys	•	1225
CTG AGT GCC ATG CGC CGC CTG GGC GTC GTG GGC GAG TTC TCA CTC ATT Leu Ser Ala Met Arg Arg Leu Gly Val Val Gly Glu Ph Ser Leu Ile 315 300 GGA AGT GAT GGA TGG GCA GAC AGA GAT GAA GTC ATC GAA GGC TAT GAG Gly Ser Asp Gly Trp Ala Asp Arg Asp Glu Val Ile Glu Gly Tyr Glu 330 330	,	GCC	C AGG	GT Va		G GTC	TGC Cys	Pne	- Cya	GAG Glu	GGG Gly	ATO			G CGC	G GGG G Gl	C TTA		1273
GGA AGT GAT GGA TGG GCA GAC AGA GAT GAA GTC ATC GAA GGC TAT GAG 1369 Gly Ser Asp Gly Trp Ala Asp Arg Asp Glu Val Ile Glu Gly Tyr Glu 330		Let	AG'		C ATO	G CGG	g Ary	nec	GGG Gly	C GTO Y Val	C GTO	,		TTO Ph	Se:	A CT r Le	C ATT u Ile 315	: *	.1321
				T GA	T GG p Gl	y Tr	b WTe	A GAC	C AG	A GA' g As	b or		C ATC	C GA	A GG	C TA y Ty 33	T GAG T Glu O	کور	1369

SUBSTITUTE SHEET

GTC GAA GCC AAC GGA GGG ATC ACA ATA AAG CTT CAG TCT CCC Val Glu Ala Asn Gly Gly Ile Thr Ile Lys Leu Gln Ser Pro 340 AGG TCA TTT GAT GAC TAC TTC CTG AAG CTG AGG CTG GAC ACC ATG Ser Phe Asp Asp Tyr Phe Leu Lys Leu Arg Leu Asp The 350 AGG AAT CCT TGG TTC CCT GAG TTC TGG CAA CAT CGC TTC CAC ATG ASN PRO Trp Phe Pro Glu Phe Trp Gln His Arg Phe Gl 370 CTA CCT GGA CAC CTC TTG GAA AAC CCC AAC TTT AAG AAA GT Leu Pro Gly His Leu Leu Glu Asn Pro Asn Phe Lys Lys Va 380 GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AAC TLY Jan Glu Ser Leu Glu Glu Asn Tyr Val Gln Asp Ser Ly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CAT CAC CAT GTG ATG ATG ATG ATG ATG ATG ATG ATG A	C AAC ACA r Asn Thr G TGT CGC n Cys Arg G TGC ACA cl Cys Thr 395 AA ATG GGA ys Met Gly 410 AG AAC ATG Ln Asn Met	1417 1465 1513 1561 1609
AGG TCA TTT GAT GAC TAC TTC CTG AAG CTG AGG CTG GAC ACC ATG SEP Phe Asp Asp Tyr Phe Leu Lys Leu Arg Leu Asp The 350 AGG AAT CCT TGG TTC CCT GAG TTC TGG CAA CAT CGC TTC CAA ARG AND Pro Trp Phe Pro Glu Phe Trp Gln His Arg Phe Gla 370 CTA CCT GGA CAC CTC TTG GAA AAC CCC AAC TTT AAG AAA GT Leu Pro Gly His Leu Leu Glu Asn Pro Asn Phe Lys Lys Va 380 GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AAC TTT GTC ATC ATC GAU Glu Glu Asn Tyr Val Gln Asp Ser Ly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CATC TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CATC TAT GCC CTC TGT GAT GCC ATG GCA CAT GGG CTG CATC TGT GAT GCC ATG GCA CAT GGG CTG CATC TGT GAT GCC ATG GCC CTC TGT GT GAT GCC ATG GCC CTC TGT GAT GCC ATG GCC CTC TGT GT GAT GCC ATG GCC CTC TGT GT GAT GCC ATG GCC CTC TGT GT GAT GCC ATG GCC ATG GCC CTC TGT GT GAT GCC ATG GCC CTC TGT GT GAT GCC ATG GCC ATG GCC ATG GCC ATG GCC CTC TGT GT GCC ATG G	C AAC ACA r Asn Thr G TGT CGC n Cys Arg G TGC ACA cl Cys Thr 395 AA ATG GGA ys Met Gly 410 AG AAC ATG Ln Asn Met	1513 1561 1609
AGG TCA TTT GAT GAC TAC TTC CTG AAG CTG AGG CTG GAC ACC ATG SEP Phe Asp Asp Tyr Phe Leu Lys Leu Arg Leu Asp The 350 AGG AAT CCT TGG TTC CCT GAG TTC TGG CAA CAT CGC TTC CAA ARG AND Pro Trp Phe Pro Glu Phe Trp Gln His Arg Phe Gla 370 CTA CCT GGA CAC CTC TTG GAA AAC CCC AAC TTT AAG AAA GT Leu Pro Gly His Leu Leu Glu Asn Pro Asn Phe Lys Lys Va 380 GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AAC TTT GTC ATC ATC GAU Glu Glu Asn Tyr Val Gln Asp Ser Ly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CATC TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CATC TAT GCC CTC TGT GAT GCC ATG GCA CAT GGG CTG CATC TGT GAT GCC ATG GCA CAT GGG CTG CATC TGT GAT GCC ATG GCC CTC TGT GT GAT GCC ATG GCC CTC TGT GAT GCC ATG GCC CTC TGT GT GAT GCC ATG GCC CTC TGT GT GAT GCC ATG GCC CTC TGT GT GAT GCC ATG GCC ATG GCC CTC TGT GT GAT GCC ATG GCC CTC TGT GT GAT GCC ATG GCC ATG GCC ATG GCC ATG GCC CTC TGT GT GCC ATG G	C AAC ACA r Asn Thr G TGT CGC n Cys Arg G TGC ACA cl Cys Thr 395 AA ATG GGA ys Met Gly 410 AG AAC ATG Ln Asn Met	1513 1561 1609
AGG TCA TTT GAT GAC TAC TTC CTG AAG CTG AGG CTG GAC ACAT GAC ATT GAT GAT GAC TAC TTC CTG AAG CTG AGG CTG GAC ACAT GAT GAT GAT CCT TAT GAT AAG AAA GT AAT GAA AAC CCC AAC TTT AAG AAA GT Leu Pro Gly His Leu Leu Glu Asn Pro Asn Phe Lys Lys Value of the Gly Fin Glu Ser Leu Glu Glu Asn Tyr Val Gln Asp Ser Ly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CATT GAT GAT GAT GAT GAT GAT GAT GAT GAT	C AAC ACA r Asn Thr G TGT CGC n Cys Arg G TGC ACA 11 Cys Thr 395 AA ATG GGA ys Met Gly 410 AG AAC ATG Ln Asn Met	1513 1561 1609
Arg Ser Phe Asp Asp Tyr Fhe 355 AGG AAT CCT TGG TTC CCT GAG TTC TGG CAA CAT CGC TTC CA Arg Asn Pro Trp Phe Pro Glu Phe Trp Gln His Arg Phe Gl 375 CTA CCT GGA CAC CTC TTG GAA AAC CCC AAC TTT AAG AAA GT Leu Pro Gly His Leu Leu Glu Asn Pro Asn Phe Lys Lys Va 385 GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AAC TTT GTC ATC ATC GLU Glu Glu Asn Tyr Val Gln Asp Ser Ly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CATC TGT GAT GAT GAT GAT GAT GAT GAT GAT GAT	TGT CGC IN CYS Arg TG TGC ACA IL CYS Thr 395 AA ATG GGA YS Met Gly 410 AG AAC ATG Ln Asn Met	1513 1561 1609
Arg Ser Phe Asp Asp Tyr Fhe 355 AGG AAT CCT TGG TTC CCT GAG TTC TGG CAA CAT CGC TTC CA Arg Asn Pro Trp Phe Pro Glu Phe Trp Gln His Arg Phe Gl 375 CTA CCT GGA CAC CTC TTG GAA AAC CCC AAC TTT AAG AAA GT Leu Pro Gly His Leu Leu Glu Asn Pro Asn Phe Lys Lys Va 385 GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AAC TTT GTC ATC ATC GLU Glu Glu Asn Tyr Val Gln Asp Ser Ly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CATC TGT GAT GAT GAT GAT GAT GAT GAT GAT GAT	TGT CGC IN CYS Arg TG TGC ACA IL CYS Thr 395 AA ATG GGA YS Met Gly 410 AG AAC ATG Ln Asn Met	1513 1561 1609
Arg Ser Phe Asp Asp Tyr Fhe 355 AGG AAT CCT TGG TTC CCT GAG TTC TGG CAA CAT CGC TTC CA Arg Asn Pro Trp Phe Pro Glu Phe Trp Gln His Arg Phe Gl 375 CTA CCT GGA CAC CTC TTG GAA AAC CCC AAC TTT AAG AAA GT Leu Pro Gly His Leu Leu Glu Asn Pro Asn Phe Lys Lys Va 385 GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AAC TTT GTC ATC ATC GLU Glu Glu Asn Tyr Val Gln Asp Ser Ly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CATC TGT GAT GAT GAT GAT GAT GAT GAT GAT GAT	TGT CGC IN CYS Arg TG TGC ACA IL CYS Thr 395 AA ATG GGA YS Met Gly 410 AG AAC ATG Ln Asn Met	1561 1609
AGG AAT CCT TGG TTC CCT GAG TTC TGG CAA CAT CGC TTC CA Arg Asn Pro Trp Phe Pro Glu Phe Trp Gln His Arg Phe Gl 375 CTA CCT GGA CAC CTC TTG GAA AAC CCC AAC TTT AAG AAA GT Leu Pro Gly His Leu Leu Glu Asn Pro Asn Phe Lys Lys Va 385 GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AAC GAY ASN GIL Fun Glu Ser Leu Glu Glu Asn Tyr Val Gln Asp Ser Ly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CAT GCC ATG GAT GAT GAT GAT GAT GAT GAT GAT GAT	TGT CGC In Cys Arg TG TGC ACA THE 395 AA ATG GGA YS Met Gly 410 AG AAC ATG Ln Asn Met	1561 1609
AGG AAT CCT TGG TTC CCT GAG TTC TGG CAA CAT CGC TTC CA Arg Asn Pro Trp Phe Pro Glu Phe Trp Gln His Arg Phe Gl 375 CTA CCT GGA CAC CTC TTG GAA AAC CCC AAC TTT AAG AAA GT Leu Pro Gly His Leu Leu Glu Asn Pro Asn Phe Lys Lys Va 380 GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AAC TTT GTC CAG GAC AGC AAC TTT GTC CAG GAC AGC AAC TTT GTC ATC ATC ATC GTU Glu Asn Tyr Val Gln Asp Ser Ly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CAT GTC ATC AAT ALA LIE TYR ALA MET ALA HIS GLY Leu GLU GLU ASN ALA LIE TYR ALA MET ALA HIS GLY LEU GLU GLU ATC TAT GCC CTC TGT GAT GAT GAT GAT GAT GAT GAT GAT GAT	TG TGC ACA 11 Cys Thr 395 AA ATG GGA 75 Met Gly 410 AG AAC ATG Ln Asn Met	1561 1609
Arg Asn Pro Trp Phe Pro Glu File 129 375 CTA CCT GGA CAC CTC TTG GAA AAC CCC AAC TTT AAG AAA GT Leu Pro Gly His Leu Leu Glu Asn Pro Asn Phe Lys Lys Va 380 GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AA Cly Fin Glu Ser Leu Glu Glu Asn Tyr Val Gln Asp Ser Ly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CA Phe Val Ile Asn Ala Ile Tyr Ala Met Ala His Gly Leu Gl 415	TG TGC ACA 11 Cys Thr 395 AA ATG GGA 75 Met Gly 410 AG AAC ATG Ln Asn Met	1561 1609
Arg Asn Pro Trp Phe Pro Glu File 129 375 CTA CCT GGA CAC CTC TTG GAA AAC CCC AAC TTT AAG AAA GT Leu Pro Gly His Leu Leu Glu Asn Pro Asn Phe Lys Lys Va 380 GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AA Cly Fin Glu Ser Leu Glu Glu Asn Tyr Val Gln Asp Ser Ly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CA Phe Val Ile Asn Ala Ile Tyr Ala Met Ala His Gly Leu Gl 415	TG TGC ACA 11 Cys Thr 395 AA ATG GGA 75 Met Gly 410 AG AAC ATG Ln Asn Met	1609
CTA CCT GGA CAC CTC TTG GAA AAC CCC AAC TTT AAG AAA GT Leu Pro Gly His Leu Leu Glu Asn Pro Asn Phe Lys Lys Va 380 GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AAC GLY Fin Glu Ser Leu Glu Glu Asn Tyr Val Gln Asp Ser Ly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CAT GCC ATG ILeu Glu Asn Ala Ile Tyr Ala Met Ala His Gly Leu Glu 420	TGC ACA 11 Cys Thr 395 AA ATG GGA 75 Met Gly 410 AG AAC ATG Ln Asn Met	1609
CTA CCT GGA CAC CTC TTG GAA AAC CCC AAC TTT AAG AAA GT Leu Pro Gly His Leu Leu Glu Asn Pro Asn Phe Lys Lys Va 380 GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AAC GLY Fin Glu Ser Leu Glu Glu Asn Tyr Val Gln Asp Ser Ly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CAT GCC ATG ILeu Glu Asn Ala Ile Tyr Ala Met Ala His Gly Leu Glu 420	TGC ACA 11 Cys Thr 395 AA ATG GGA 75 Met Gly 410 AG AAC ATG Ln Asn Met	1609
CTA CCT GGA CAC CTC TTG GAA AAC CCC AAC TTT AAG AAA GT Leu Pro Gly His Leu Leu Glu Asn Pro Asn Phe Lys Lys Va 380 GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AAC GLY Fin Glu Ser Leu Glu Glu Asn Tyr Val Gln Asp Ser Ly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CAT Phe Val Ile Asn Ala Ile Tyr Ala Met Ala His Gly Leu Gly 420	395 AA ATG GGA AS Met Gly 410 AG AAC ATG Ln Asn Met	1609
Leu Pro Gly His Leu Leu Glu Ash 710 390 GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AA Cly Fin Glu Ser Leu Glu Glu Ash Tyr Val Gln Asp Ser Ly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CA Phe Val Ile Ash Ala Ile Tyr Ala Met Ala His Gly Leu Gl 420 420	395 AA ATG GGA AS Met Gly 410 AG AAC ATG Ln Asn Met	
Leu Pro Gly His Leu Leu Glu Ash 710 390 GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AA Cly Fin Glu Ser Leu Glu Glu Ash Tyr Val Gln Asp Ser Ly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CA Phe Val Ile Ash Ala Ile Tyr Ala Met Ala His Gly Leu Gl 420 420	395 AA ATG GGA AS Met Gly 410 AG AAC ATG Ln Asn Met	
GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AAC GLY Fin Glu Ser Leu Glu Glu Asn Tyr Val Gln Asp Ser Ly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CAT Phe Val Ile Asn Ala Ile Tyr Ala Met Ala His Gly Leu Gly 420	AA ATG GGA ys Met Gly 410 AG AAC ATG Ln Asn Met	
GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AA Cly Fin Glu Ser Leu Glu Glu Asn Tyr Val Gln Asp Ser Ly 400 TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CA Phe Val Ile Asn Ala Ile Tyr Ala Met Ala His Gly Leu Gl 420 420 TTT GTC GGC CTC TGT GAT GCC	410 AG AAC ATG In Asn Met 25	
TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CAT Phe Val Ile Asn Ala Ile Tyr Ala Met Ala His Gly Leu Gl 420	410 AG AAC ATG In Asn Met 25	
TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CAT Phe Val Ile Asn Ala Ile Tyr Ala Met Ala His Gly Leu Gl 420	410 AG AAC ATG In Asn Met 25	
TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CAT Phe Val Ile Asn Ala Ile Tyr Ala Met Ala His Gly Leu Gl 420	AG AAC ATG Ln Asn Met 25	1657
TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CAT Phe Val Ile Asn Ala Ile Tyr Ala Met Ala His Gly Leu Gl 420	25	1657
Phe Val Ile Ash Ala Ile Tyl Ala Med Med 420 415 420 420 427 420 420	25	1657
Phe Val Ile Ash Ala Ile Tyl Ala Med Med 420 415 420 420 427 420 420	25	•
Phe Val Ile Ash Ala Ile Tyl Ala Med Med 420 415 420 420 427 420 420	25	
415		• •
THE SEC CAM STG CGC CTG TGT GAT GC	1 1	·
AND COLOR OF CALL CALL CALL CALL CALL CALL CALL CAL	ግጥ አጥር <u>አ</u> ልሕ	1705
ON CAM COM CITE TET CCC GGC CAI GIG GGC CAI	La Mot IVE	
CAC CAT GCT CTG TGT CCC GGC CAT GTG GGC CTG TGT CALL His His Ala Leu Cys Pro Gly His Val Gly Leu Cys Asp Al 440	Ta wer Dia	•
HIS HIS AIR LEG CID LEG 435	*	•
430		1767
CCC ATT GAT GGC AGG AAG CTC CTG GAT TTC CTC ATC AAA TO	CC TCT TTT	1753
CCC ATT GAT GGC AGG AAG CIC CIG AGD Phe Leu Ile Lys So	er Ser Phe	
Dro Tle ASD Gly Arg Lys Leu Leu Moy 188	•	
445	• • •	
GTC GGA GTG TCT GGA GAG GAG GTG TGG TTC GAT GAG AAG GG	GG GAT GCT	1801
CTC GGA GTG TCT GGA GAG GAG GTG TGG TTC GAT GAG AAG G	ly Asp Ala	÷
" wal gla Val SAT GIV GIU GIU VAL INF	475	
465	7,0	
400		1849
CCC GGA AGG TAT GAC ATT ATG AAT CTG CAG TAC ACA GAA G	CT AAT CGC	
CCC GGA AGG TAT GAC ALL ALL ALL Gln Tyr Thr Glu A	la Asn Arg	
Pro Gly Arg Tyr Asp Tie Met Ash Age	490	
400		
TAT GAC TAT GTC CAC GTG GGG ACC TGG CAT GAA GGA GTG C	TG AAT ATT	1897
TAT GAC TAT GTC CAC GTG GGG ACC TGG CAT GAA GGN Val L Tyr Asp Tyr Val His Val Gly Thr Trp His Glu Gly Val L 500	eu Asn Ile	
men han mur Val HIS Val GLY INA 1-5	05	
19F ASP 191 495 500		
		1945
GAT GAT TAC AAA ATC CAG ATG AAC AAA AGC GGA ATG GTA C	GA TOT GIG	
GAT GAT TAC AMA ALC CAS Not Ash Lys Ser Gly Met Val A	rd Ser Agr	
GAT GAT TAC AAA ATC CAG ATG AAC AAA AGC GGA ATC AAA AAA ACC AAA AAA ACC AAA AAA AA		
510	•	
TGC AGT GAG CCT TGC TTA AAG GGT CAG ATT AAG GTC ATA C	GG AAA GGA	1993
TCC AGT GAG CCT TGC TTA AAG GGT CAG ATT AAG US TILD A	ra Lvs Gly	7
TGC AGT GAG CCT TGC TTA AAG GGT CAG ATT AAG STO ATT TGC AGT GAG CCT TGC TTA AAG GGT CAG ATT AAG STO ATT TGC AGT CAG ATT AAG STO AGT AAG ATT AAG ST		· . · · · ·
535 525		
	Am Cac men	2041
GAA GTG AGC TGC TGC ATC TGC ACG GCC TGC AAA GAG A	MI GRG III	
GAA GTG AGC TGC TGC ATC TGC ACG GCC 1GC ALL GIU A Glu Val Ser Cys Cys Trp Ile Cys Thr Ala Cys Lys Glu A 550	ARM GIR LUG	
Glu Val Ser Cys Cys Trp 118 Of 550	555	
540		•
THE SEA COO MOTE CAN CTIC GGG T	rgg tgg ccc	2089
GTG CAG GAC GAG TTC ACC TGC AGA GCC TGT GAC CTG GGG T		
	•	
SUBSTITUTE SHEET	•	
	Ť	

Val Gln Asp Glu Phe Thr Cys Arg Ala Cys Asp Leu Gly Trp Try 560 565 576 576 576 576 576 565 576 576 576	r GAG 2137 u Glu G GGC 2185 u Gly G GAC 2233
Asn Ala Glu Leu Thr Gly Cys Glu Pro 11e Pro 585 TGG AGT GAC ATA GAA TCT ATC ATA GCC ATC GCC TTT TCT TGC CTC Trp Ser Asp Ile Glu Ser Ile Ile Ala Ile Ala Phe Ser Cys Le 590 ATC CTC GTG ACG CTG TTT GTC ACC CTC ATC TTC GTT CTG TAC CG Ile Leu Val Thr Leu Phe Val Thr Leu Ile Phe Val Leu Tyr Ar	G GGC 2185 u Gly G GAC 2233
TGG AGT GAC ATA GAA TCT ATC ATA GCC ATC GCC TTT TCT TGC CTC TTP Ser Asp Ile Glu Ser Ile Ile Ala Ile Ala Phe Ser Cys Les 590 ATC CTC GTG ACG CTG TTT GTC ACC CTC ATC TTC GTT CTG TAC CG Ile Leu Val Thr Leu Phe Val Thr Leu Ile Phe Val Leu Tyr Ar	G GAC 2233
Trp Ser Asp Ile Glu Ser Ile Ile Ald Ile Ald 600 595 ATC CTC GTG ACG CTG TTT GTC ACC CTC ATC TTC GTT CTG TAC CG Ile Leu Val Thr Leu Phe Val Thr Leu Ile Phe Val Leu Tyr Ar	G GAC 2233
Ile Leu Val Thr Leu Phe Val Thr Leu Ile Phe Val Bed 1/2	g GAC 2233 g Asp
615	
ACA CCC GTG GTC AAA TCC TCC AGT AGG GAG CTC TGC TAT ATC AT Thr Pro Val Val Lys Ser Ser Ser Arg Glu Leu Cys Tyr Ile Il 620	T CTG 2281 e Leu 635
GCT GGT ATT TTC CTC GGC TAT GTG TGC CCT TTC ACC CTC ATC GC Ala Gly Ile Phe Leu Gly Tyr Val Cys Pro Phe Thr Leu Ile Al 640	C AAA 2329 a Lys
CCT ACT ACC ACA TCC TGC TAC CTC CAG CGC CTC CTA GTT GGC CT Pro Thr Thr Ser Cys Tyr Leu Gln Arg Leu Leu Val Gly Le 655	C TCT 2377 au Ser
TCT GCC ATG TGC TAC TCT GCT TTA GTG ACC AAA ACC AAT CGT AT Ser Ala Met Cys Tyr Ser Ala Leu Val Thr Lys Thr Asn Arg II 670 675	TT GCA 2425 Le Ala
CGC ATC CTG GCT GGC AGC AAG AAG ATC TGC ACC CGG AAG CC Arg Ile Leu Ala Gly Ser Lys Lys Ile Cys Thr Arg Lys Pi 685 690	CC AGA 2473 co Arg
TTC ATG AGC GCT TGG GCC CAA GTG ATC ATA GCC TCC ATT CTG ATC ATC ATG AGC Phe Met Ser Ala Trp Ala Gln Val Ile Ile Ala Ser Ile Leu I. 700 705	TT AGT 2521 le Ser 715
GTA CAG CTA ACA CTA GTG GTG ACC TTG ATC ATC ATG GAG CCT C	CC ATG 2569 ro Met 30
CCC ATT TTG TCC TAC CCG AGT ATC AAG GAA GTC TAC CTT ATC T Pro Ile Leu Ser Tyr Pro Ser Ile Lys Glu Val Tyr Leu Ile C 735	GC AAT 2617 ys Asn
ACC AGC AAC CTG GGT GTA GTG GCC CCT GTG GGT TAC AAT GGA C Thr Ser Asn Leu Gly Val Val Ala Pr Val Gly Tyr Asn Gly L 750 760	TC CTC 2665 eu Leu
ATC ATG AGC TGT ACC TAC TAT GCC TTC AAG ACC CGC AAC GTG C Ile Met Ser Cys Thr Tyr Tyr Ala Phe Lys Thr Arg Asn Val P	CG GCC 2713
AAC TTC AAT GAG GCT AAA TAC ATC GCC TTC ACC ATG TAC ACT A Asn Phe Asn Glu Ala Lys Tyr Ile Ala Ph Thr Met Tyr Thr T	cc TGC 2761 hr Cys

7	80		• **			785			,		790		٠,	•		795		C = 1
]	TC le	ATC Ile	TGG	CTG Leu	GCT Ala 800	Phe	GTT Val	CCC	ATT	TAC Tyr 805	TTT Phe	GGG Gly	AGC Ser	AAC Asn	TAC Tyr 810	AAG Lys		2809
Ā	TC le	ATC	ACI	ACC Thr 815	Cys	TTC Phe	GCG Ala	GTG Val	AGC Ser 820	Leu	AGT Ser	GTG Val	ACG Thr	GTG Val 825	Ala	CTG Leu		2857
G	GG ly	TGC Cys	ATG Met 830	Phe	ACT	CCG Pro	AAG Lys	ATG Met 835	TAC	ATC Ile	ATC	Ile	GCC Ala 840	AAA Lys	CCT	GAG Glu	•	2905
A	GG rg	AAC Asn 845	Val	CGC Arg	AGT Ser	GCC	TTC Phe 850	Thr	ACC	TCT Ser	GAT Asp	GTT Val 855	Val	CGC	ATG Met	CAC His	,	2953
V	TC al 60	G GT Gly	GAT	GGC	AAA Lys	CTG Leu 865	CCG Pro	TGC Cys	CGC	TCC Ser	AAC Asn 870	ACC	TTC Phe	CTC	AAC Asn	ATT Ile 875		3001
T	TC he	CGG Arg	AGA Arg	AAG Lys	AAG Lys 880	CCC	GGG Gly	GCA Ala	GGG	AAT Asn 885	GCC Ala	AAT Asn	TCT	AAC Asn	GGC Gly 890	AAG Lys	· · · · ·	3049
T	CT er	GTG Val	TCA Ser	TGG Trp 895	TCT Ser	GAA Glu	CCA Pro	GGT Gly	GGA Gly 900	Arg	CAG Gln	GCG Ala	CCC	AAG Lys 905	Gly	CAG Gln		3097
C H	AC is	GTG Val	TGG Trp 910	CAG Gln	CGC	CTC	TCT	GTG Val 915	CAC His	GTG Val	AAG Lys	ACC	AAC Asn 920	GAG Glu	ACG	GCC Ala		3145
T	ys	AAC Asn 925	Gln	ACA Thr	GCC Ala	Val	ATC Ile 930	Lys	CCC	CTC	ACT Thr	AAA Lys 935	AGT Ser	TAC Tyr	CAA Gln	GGC Gly		3193
S	CT er 40	GGC Gly	AAG Lys	AGC Ser	CTG Leu	ACC Thr 945	TTT Phe	TCA Ser	GAT Asp	GCC Ala	AGC Ser 950	ACC	AAG Lys	ACC Thr	CTT Leu	TAC Tyr 955		3241
A.	AT sn	GTG Val	GAA Glu	GAA Glu	GAG Glu 960	GAC Asp	AAT Asn	ACC	CCT	TCT Ser 965	GCT Ala	CAC His	TTC Phe	AGC Ser	CCT Pro 970	CCC		3289
A(GC er	AGC Ser	CCT	TCT Ser 975	ATG Met	GTG Val	GTG Val	CAC His	CGA Arg 980	CGC	GGG	CCA Pro	CCC Pro	GTG Val 985	GCC Ala	ACC Thr		3337
A(ZA I	Pro	CCT Pro 990	CTG Leu	CCA	CCC Pro	His	CTG Leu 995	ACC Thr	GCA Ala	GAA Glu	GAG Glu	ACC Thr 1000	PTO	CTG	TTC Phe		3385
CI	n i	GCT Ala 1005	qaƙ	TCC Ser	GTC Val	Ile	CCC Pro 1010	Lys	GGC Gly	TTG Leu	CCT Pr	CCT Pr 1015	CCT Pr	CTC Leu	CCG Pro	CAG Gln		3433

CAG-CAG CCA CAG CAG CCC CCT CAG CAG CCC CCG CAG CAG CCC AAG Gln Gln Pro Gln Gln Pro Pro Gln Gln Pro Pro Gln Gln Pro Lys 1020 1035	3481
TCC CTG ATG GAC CAG CTG CAA GGC GTA GTC ACC AAC TTC GGT TCG GGG Ser Leu Met Asp Gln Leu Gln Gly Val Val Thr Asn Phe Gly Ser Gly 1040	3529
ATT CCA GAT TTC CAT GCG GTG CTG GCA GGC CCG GGG ACA CCA GGA AAC Ile Pro Asp Phe His Ala Val Leu Ala Gly Pro Gly Thr Pro Gly Asn 1055	3577
AGC CTG CGC TCT CTG TAC CCG CCC CCG CCT CCG CCG CAA CAC CTG CAG Ser Leu Arg Ser Leu Tyr Pro Pro Pro Pro Pro Pro Gln His Leu Gln 1070	3625
ATG CTG CCC CTG CAC CTG AGC ACC TTC CAG GAG GAG TCC ATC TCC CCT Met I.au Pro Leu His Leu Ser Thr Phe Gln Glu Glu Ser Ile Ser Pro 1085	3673
CCT GGG GAG GAC ATC GAT GAC AGT GAG AGA TTC AAG CTC CTG CAG Pro Gly Glu Asp Ile Asp Asp Asp Ser Glu Arg Phe Lys Leu Leu Gln 1100 1115	3721
GAG TTC GTG TAC GAG CGC GAA GGG AAC ACC GAA GAA GAT GAA TTG GAA GIU Phe Val Tyr Glu Arg Glu Gly Asn Thr Glu Glu Asp Glu Leu Glu 1120	3769
GAG GAG GAC CTG CCC ACA GCC AGC AAG CTG ACC CCT GAG GAT TCT Glu Glu Glu Asp Leu Pro Thr Ala Ser Lys Leu Thr Pro Glu Asp Ser 1135	3817
CCT GCC CTG ACG CCT CCT TCT CCT TTC CGA GAT TCC GTG GCC TCT GGC Pro Ala Leu Thr Pro Pro Ser Pro Phe Arg Asp Ser Val Ala Ser Gly 1150 1155	3865
AGC TCA GTG CCC AGT TCC CCC GTA TCT GAG TCG GTC CTC TGC ACC CCT Ser Ser Val Pro Ser Ser Pro Val Ser Glu Ser Val Leu Cys Thr Pro 1165	3913
CCA AAT GTA ACC TAC GCC TCT GTC ATT CTG AGG GAC TAC AAG CAA AGC Pro Asn Val Thr Tyr Ala Ser Val Ile Leu Arg Asp Tyr Lys Gln Ser 1180 1195	3961
TCT TCC ACC CTG TAGTGTGTGT GTGTGTGTGG GGGCGGGGG AGTGCGCATG Ser Ser Thr Leu	4013
GAGAAGCCAG AGATGCCAAG GAGTGTCAAC CCTTCCAGAA ATGTGTAGAA AGCAGGGTGA	4073
GGGATGGGGA TGGAGGACCA CGGTCTGCAG GGAAGAAAA AAAAATGCTG CGGCTGCCTT	4133
AAAGAAGGAG AGGGACGATG CCAACTGAAC AGTGGTCCTG GCCAGGATTG TGACTCTTGA	4193
ATTATTCAAA AACCTTCTCT AGAAAGAAAG GGAATTATGA CAAAGCACAA TTCCATATGG	4253
TATGTAACIT TTATCGAAAA AAAAAAAAAA AAAAAAAAA AAAAAAA	4300

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1199 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

	•	-	اعبدن			• •		. *		•					•	
٠,	(x:	i) s	EQUE	NCE	DESC	RIPT	ION:	SEQ	ID	NO:2	:		* * * * * * * * * * * * * * * * * * * *			•
Met \	Val	Arg	Leu	Leu 5	Leu	Ile	Phe	Phe	Pro 10	Met	Ile					
Ser	٠.		ZU				-									
Ser		35						•								
Ala	らり		·	•		J.J	•						•			
~ 	Lys			- ,	Ile 70											0
Ala	Met	Phe	His	Thr 85	Leu	Asp	Lys	Ile	Asn 90	Ala	Asp	Pro	Val	Leu 95	Leu	-
Pro	Asn	Ile	Thr 100		Gly	Ser	Glu	Ile 105	Arg	Asp	Ser	Cys	Trp 110	His	Ser	*
Ser	Val	Ala 115	Leu	Glu	Gln	Ser	Ile 120	Glu	Phe	Ile	Arg	Asp 125	Ser	Leu	Ile	,
Ser	Ile	Arg	Asp	Glu	Lys	Asp 135	Gly	Leu	Asn	Arg	Cys 140	Leu	Pro	Asp	Gly	,
Gln	130 Thr	Leu	Pro	Pro	Gly 150	Arg		. Lys	Lys	Pro 155	Ile	Ala	Gly	val	Ile 160	
145 Gly	Pro	Gly	ser	Ser 165	: Ser		Ala	Ile	Gln 170	val	Gln	Asn	Leu	Leu 175	Gln	
Leu	Phe	ÀSŢ	Ile	Pro	Gln	Ile	Ala	Tyr 185	Ser	: Ala	Thr	Ser	190	a Asp	Leu	
Ser	Asp	Lys	180 Thi	Let	ı Tyr	Lys	TY1	Phe		ı Arg	Val	Va]	Pro	Ser	Asp	. '
		14.	•				t Le					•			Trp	
	210 Tyr					L Hi	_								c Gly 240	

			•			;	14			, ,	•					
Met	As	sp i	Ala	Phe	Lys 245	Glu	Leu	Ala	Ala	Gln 250	Glu	Gly	Leu	Cys	11e 2 255	Ala
		•		260	"	-			205	•				Phe 270		
Leu	L	eu	Arg 275	Lys	Leu	Arg	Glu	Arg 280	Leu	Pro	Lys	Ala	Arg 285	Val	Val	Val
•	2	90		;	٠,.		295			0.0				Ala	•	
305	5		*			310					323			Asp		
					325		•	•		330	•		•	Ala	•	
,				340				•	, 3,43	•			. ,	Phe 350		•
•	,		355		•			300		• • •				Pro	٠.,	*
	. 3	70		•	ī	o e	3/5	•		\$ 1				Gly		
38	5		-			390			•	_		•		Glu		
					405	•				720	•			Ile		
-		•		420	ָי				, 72-	* .			•	e ·	-	Cys
		٠.	435		• •		-	441	•				9		6	Arg
		450					45	•			*	•	•			Gly
46	5		-	į	84	470	J	· . •	•	• -					•	Asp 480
,			·	1	48	•	•			*		•			:	His
		٠	•	50	0		· •	,	. 50	-	* . *	. *		8		Ile
G.	Ln	Met	: As: 51	n Ly 5	s Se		-	52	J							cys
L	eu ,	Lys 530		y Gl	n Il	Ly	s Va 53		e Ar	g Ly	s Gl	y G1 54	u Va	l Sei	Cyt	Cys

•					•. •			* ,		,		•					
	545			,	•	. 0	950								Asp	•	
						202	•								Glu		
					58 U					-					Asp 590	•	
			59	5					000				•		Val		•
		610			-			013	-						Val		٠.
,	Ser 625	Ser	Se	er	Arg	Glu	Leu 630	Cys	Tyr	Ile	Ile	Leu 635	Ala	Gly	Ile	Phe	Leu 640
	Gly	туг	· Va	ıl	Cys	Pro 645	Phe	Thr	Leu	Ile	Ala 650	Lys	Pro	Thr	Thr	Thr 655	Ser
	Cys	Tyr	Le	₃u'	Gln 660	Arg	Leu	Leu	Val	Gly 665	Leu	Ser	Ser	Ala	Met 670	Cys	Tyr
	Ser	Ala	Le 67	au 75	Val	Thr	Lys	Thr	Asn 680	Arg	Ile	Ala	Arg	11e 685	Leu	Ala	Gly
	Ser	Ly:	s Ly			Ile	Cys	Thr 695	Arg	Lys	Pro	Arg	Phe 700	Met	Ser	Ala	Trp
	Ala 705	Gl	n V	al	Ile	Ile	Ala 710	Ser	: Ile	Leu	Ile	Ser 715	Val	Gln	Leu	Thr	Leu 720
	Val	. Va	l T	hr	Leu	11e	ile	Met	e Glu	Pro	730	Met	Pro	Ile	Leu	Ser 735	Tyr
	Pro	Se	r I	le	Lys	Glu	ı Val	L Tyi	r Lev	1 Ile 745	Cys	a Asn	Thr	Ser	750	Leu	Gly
	. Val	L Va	1 A	1a 55	Pro		L Gly	у Ту	r Asr 760	Gly	, Lev	ı Lev	ı Ile	Met 765	: Ser	Cys	Thr
	Tyı	r Ty 77	r A	la	Phe	_	s Th	77	g Ası 5	va]	l Pro	Ala	780	n Phe	a Asn	Glu	Ala
		з Ту	r I	:le	Ala	a Ph	e Th:	r Me	t Ty	r Thi	r Thi	795	s Ile 5	a Ile	e Trp	Lev	Ala 800
	789 Phe	e Va	1 F	ro	Ile	e Ty:	r Ph	•	y Se	r Asi	n Ty:	r Ly:	s Ile	e Ile	e Thi	Thr 815	Cys
	Pho	e Al	a V	7al	. Se:	r Le		r Va	1 Th:	r Va 82	l Al	a Lei	u Gly	y Cy	s Met	t Ph€	Thr
	, Pr	o Ly	s l	let 335	Ty	r Il	e Il	e Il	e Al 84	a Ly O	s Pr	Gl	u Ar	g As 84	n Va: 5	l Arq	g Ser
			•		-										•		

	. *				4			•		,					_ :
	850	• • •			1	855		•			500			Gly	•
865					87.0		•			6/5	-		•		
Pro	Gly	Ala	Gly	Asn 885	Ala	Asn	Ser	Asn	Gly 890	Lys	Ser	Val	Ser	Trp 895	Ser
Glu	Pro	Gly	Gly 900	Arg	Gln	Ala	Pro	Lys 905	Gly	Glņ	His	Val	Trp 910	Gln	Arg
Leu	Ser	Val 915	His	Val	Lys	Thr	Asn 920	Glu	Thr	Ala	Cys	Asn 925	Gln	Thr	Ala
Val	Ile 930	Lys	Pro	Leu	Thr	Lys 935	Ser	Tyr	Gln	Gly	Ser 940	Gly	Lys	Ser	Leu
Thr 945	Phe	Ser	Asp	Ala	Ser 950	Thr	Lys	Thr	Leu	Tyr 955	Asn	Val	Glu	Glu	Glu 960
Asp	Asn	Thr	Pro	Ser 965	Ala	His	Phe	Ser	Pro 970	Pro	Ser	Ser	Pro	Ser 975	Met
Val	Väl	His	Arg 980	Arg	Gly	Pro	Pro	Val 985	Ala	Thr	Thr	Pro	Pro 990	Leu	Pro
Pro	His	Leu 995		Ala	Glu	Glu	Thr 100	Pro 0	Leu	Phe	Leu	Ala 100	Asp 5	Ser	Val
Ile	Pro 101		Gly	Leu	Pro	Pro 101	Pro 5	Leu	Pro	Gln	Gln 102	Gln O	Pro	Gln	Gln
Pro		Pro	Gln	Gln	Pro 103	Pro	Gln	Gln	Pro	Lys 103	Ser 5	Leu	Met	Asp	Gln 1046
Leu	Gln	Gly	val	'Val	Thr 5	Asn	Phe	Gly	Ser 105	Gly O	: Ile	Pro	Asp	Phe 105	His 5
Ala	Val	Lev	Ala 105	Gly	Pro	Gly	Thr	Pro 106	Gly 5	Asn	Ser	Leu	Arg 107	Ser O	Leu
Туг	Pro	Pro 107		Pro	Pro	Pro	Gln 108	His O	Leu	Gln	Met	Leu 108	Pro	Leu	His
Leu	Ser 109	Thi	r Phe	Gln	Glu	Glu 109	Ser 5	Ile	Ser	Pro	Pro 110	Gly	Glu	Asp	Ile
Asp 110) Ası		Ser	: Glu	Arg	Phe 0	Lys	Leu	Lev	Glr	Glu 15	Phe	val	Tyr	Glu 112
		ı Gl	y Asi	Th:	- Glu 25	Glu	ı Asp	Glu	Leu 113	ı Glu 30	i Glu	ı Glu	ı Glu	1 Asp 113	Leu 5
Pr	Thi	r Ala	a Sei 114		Leu			114	כו		e Pro	Ala	11!	Thi	Pr
**						SU	BSTI	TUTI	E SH	EET				•	

Pro Ser Pro Phe Arg Asp Ser Val Ala Ser Gly Ser Ser Val Pro Ser 1155

Ser Pro Val Ser Glu Ser Val Leu Cys Thr Pro Pro Asn Val Thr Tyr 1170 1180

Ala Ser Val Ile Leu Arg Asp Tyr Lys Gln Ser Ser Ser Thr Leu 1185 1190 1195

- (2) INFORMATION FOR SEQ ID NO:3:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - 'ii) MOLECULE TYPE: cDNA
 - (vii) IMMEDIATE SOURCE: (B) CLONE: ZC775
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3: -

GCTAGCATAA CCCCTTGGGG CCTCTAAACG GGTCT

- (2) INFORMATION FOR SEQ ID NO:4:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 43 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: CDNA
 - (vii) IMMEDIATE SOURCE:
 - (B) CLONE: ZC776
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CTCAAGACCC GTTTAGAGGC CCCAAGGGGT TATGCTAGCT GCA

- (2) INFORMATION FOR SEQ ID NO:5:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 40 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: lin ar
 - (ii) MOLECULE TYPE: CDNA

SUBSTITUTE SHEET

(Vii)	IMMEI (B)	CLONE	SOUR ZC	777	
•		· -			
(xi)	SEQU	ENCE D	ESCR	IPTIO	N:

(XI) SEQUENCE DESCRIPTION: SEQ ID NO:5:
TGAGGGGTTT TTTGCTGAAA GGAGGAACTA TGCGGCCGCA

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 40 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(Vii) IMMEDIATE SOURCE: (B) CLONE: ZC778

(X1) SEQUENCE DESCRIPTION: SEQ ID NO:6: AGCTTGCGGC CGCATAGTTC CTCCTTTCAG CAAAAAACCC

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: CDNA

(vii) IMMEDIATE SOURCE:

(B) CLONE: ZC1751

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AATTCTGTGC TCTGTCAAG

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: singl

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: CDNA

40

40

-(vii) IMMEDIATE SOURCE: (B) CLONE: ZC1752

(XI) SEQUENCE DESCRIPTION: SEQ ID NO:8: GATCCTTGAC AGAGCACAG

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: CDNA

(vii) IMMEDIATE SOURCE: (B) CLONE: ZC2063

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GATCCAAACT AGTAAAAGAG CT

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 base pairs (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: CDNA

(Vii) IMMEDIATE SOURCE: (B) CLONE: ZC2064

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CTTTTACTAG TTTG

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: CDNA

22

14

19

CHRSTITUTE SHEET

(Vii) IMMEDIATE SOURCE:
(B) CLONE: ZC2938

(XI) SEQUENCE DESCRIPTION: SEQ ID NO:11: GACAGAGCAC AGATTCACTA GTGAGCTCTT TTTTTTTTT TTT

43

- (2) INFORMATION FOR SEQ ID NO:12:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: CDNA
 - (Vii) IMMEDIATE SOURCE: (B) CLONE: ZC3015
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TTCCATGGCA CCGTCAAGGC T

21

- (2) INFORMATION FOR SEQ ID NO:13:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: CDNA
 - (vii) IMMEDIATE SOURCE: (B) CLONE: ZC3016
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AGTGATGGCA TGGACTGTGG T

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucl ic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: CDNA

	8 ·	2.		•	FF
(vii) IMMEDIATE	SOURCE:				· .
(B) CLONE	2: ZC3652	4 · * 0		•	
	•		•	•	8
(xi) SEQUENCE I	DESCRIPTION: SEQ	ID NO:14:			20
ACATGCACCA TGCTCTG	rgt	· · · · · · · · · · · · · · · · · · ·	* *		
(2) INFORMATION FOR	R SEQ ID NO:15:		*		•
(A) LENG (B) TYPE (C) STRA	CHARACTERISTICS: TH: 21 base pair: : nucleic acid NDEDNESS: single LOGY: linear	-		· · · · · · · · · · · · · · · · · · ·	
(ii) MOLECULE	TYPE: CDNA				-
	*	•			
			0	*	•
(vii) IMMEDIATE	SOURCE:		. * *		- × .
(B) CION	E: ZC3654	-			
			•		8
(xi) SEQUENCE	DESCRIPTION: SEQ	ID NO:15:			21
AGTGATGGCA TGGACTG			**		
(2) INFORMATION FO	R SEQ ID NO:16:				
(A) LENG (B) TYPE (C) STRE	CHARACTERISTICS: TH: 5236 base pa : nucleic acid NDEDNESS: single NLOGY: linear	TLS			. •)•
(ii) MOLECULE	TYPE: CDNA				•
	- COURCE:	,	*		•
(vii) IMMEDIATI (B) CLOR	NE: SN23			0	
	**	H	*		•
(ix) FEATURE: (A) NAM	E/KEY: CDS ATION: 6273344	*			• • .
(B) INC	WIIOM. OBLITAGE	*			•
	·		•	*	
(xi) SEQUENCE	DESCRIPTION: SEC	J ID MO:16:		TA CA GGGCAC	60
TAAGAATTIT ATAAAT	ACTC TGGGAATTTT	ATTGGTGATG	CCLLEGIGIC	1 COMMUNICA A	120
ACGTTCCAGA GAGCTC	TGGT GTGAAGTGAT	GGGGGACTTG	TGGCTAGAGA	AGCTTTTCAA	•
TGGCCTTAAA CTCTGG	GTCC TGCTTGAGAG	AGGTCTGAGG	TTCTCAACAT	CAGAGCAGAG	240
CTTCCACCAA GCTTTC	AGAA TGCTAAGCCC	CCACTTCTCA	ACACTTAGTG	CTCTGATCGG	300
TGCCTGCGAA CCGAGA	ACGG CTGCAGTCCT	CTGACCTGAG	ACCAATAGCT	GTGTCTACCC	- 300

	360
GGACTCAGCG TCCAGCTCAC CGCCACTAAC GCGCCGCGCA TTGGACACCT GATCCACACA	700
CCTTCGGGCA CCAGTGAAAA ACCGCGACTT GATTTTCTGG AAGAACGCCC CCAGGGTGTG	420
GGAGCGGTCG TGGAGGACCA GCAGGAGGAA GCGGAGGGGA GAGGGGCAGT AGTGGAGGCA	480
GAGAAAGCGT TGAACCAGCT GTGTTGGCCG AAGGCACGAA ACGGCAAAAG GCAGCGGTGA	540
	600
GCATCTGTGT GGTTCCCGCT GGGAACCTGC AGGCAGGACC GGCGTGGGAA CGTGGCTGGC	800
CCGCGGTGGA CCGCGTCTTC GCCACA ATG GTC CGG CTC CTC TTG ATT TTC TTC Met Val Arg Leu Leu Ile Phe Phe 5	653
CCA ATG ATC TTT TTG GAG ATG TCC ATT TTG CCC AGG ATG CCT GAC AGA	701
Pro Met Ile Phe Leu Glu Met Ser Ile Leu Pro Arg Met Plo Asp 25	•
10	749
AAA GTA TTG CTG GCA GGT GCC TCG TCC CAG CGC TCC GTG GCG AGA ATG	143
Lys Val Leu Leu Ala Gly Ala Ser Ser Gln Arg Ser Val Ala Arg Met 30 35	
	797
GAC GGA GAT GTC ATC GGA GCC CTC TTC TCA GTC CAT CAC CAG CCT Asp Gly Asp Val Ile Ile Gly Ala Leu Phe Ser Val His His Gln Pro	1.
Asp GIV Asp val Tie Tie GIV AIG 55	
THE STATE OF THE SEC CAG ATC AGG GAA CAG	845
CCA GCC GAG AAG GTA CCC GAA AGG AAG TGT GGG GAG ATC AGG GAA CAG Pro Ala Glu Lys Val Pro Glu Arg Lys Cys Gly Glu Ile Arg Glu Gln	
60 65 70	
TAT GGT ATC CAG AGG GTG GAG GCC ATG TTC CAC ACG TTG GAT AAG ATT	893
Tur Gly Ile Gln Ard Val Glu Ala Met Phe His ini bed Asp 2/3 220	
75 80	
AAC GCG GAC CCG GTG CTC CTG CCC AAC ATC ACT CTG GGC AGT GAG ATC	941
len Ala Asp Pro Val Leu Leu Pro Ash Tie Thr Leu Giy Ser Ser 105	•
90 95	
CGG GAC TCC TGC TGG CAC TCT TCA GTG GCT CTC GAA CAG AGC ATC GAA	989
Arg Asp Ser Cys Trp His Ser Ser val Ala Leu Glu Glu Glu Glu Glu Glu Glu Glu Glu Gl	
2. 110	
TTC ATC AGA GAC TCC CTG ATT TCC ATC CGA GAT GAG AAG GAT GGG CTG	1037
Phe Ile Arg Asp Ser Leu Ile Ser Ile Arg Asp Git Lys Asp Gill Lys Asp G	· · · · · ·
	1005
AAC CGA TGC CTG CCT GAT GGC CAG ACC CTG CCC CCT GGC AGG ACT AAG	1085
Asn Arg Cys Leu Pro Asp Gly Gin Thr Leu Pro Pro Gly Alg	
140	1133
AAG CCT ATT GCT GGA GTG ATC GGC CCT GGC TCC AGC TCT GTG GCC ATT	
Lys Pr Ile Ala Gly val II Gly Plo Gly Del 165	
	1181
CAA GTC CAG AAT CTT CTC CAG CTG TTC GAC ATC CCA CAG ATC GCC TAT Gln Val Gln Asn Leu Leu Gln Leu Phe Asp Ile Pr Gln Il Ala Tyr 180	
Gln Val Gln Asn Leu Lau Gln Leu File Asp 120 185	

			•	1, 10		•		ē .	84		0	7	18			σ.	
*				1							man er	መአድ	777	TAC	TTC	. 12	29
-ئىسان	GCC	ACA	AGC	ATA-	GAC	CTG	AGT .	GAC	AAA	ACT	TIG	TUT	ann Tue	TVT.	Phe	7.0	10
CAT	Ala	Thr	Ser	Ile	Asp.	Leu	Ser	Asp	TAR	THE	Leu	- 7 -	2,2	200	10		1.5
501 .		-		190					Tab	,	τ.,				, , , , , , , , , , , , , , , , , , ,		
									oic.	CCA	AGG	GCG	ATG	CTC	GAC	12	277
CTG	AGG	GTG	GTC	CCT	TCT	GAC	ACT	TTG	CAG	Ala	ATC	Ala	Met	Leu	Asp		
Leu	Arg	Val	Val	Pro	Ser	GAC Asp	Thr	Leu	GTU	WIG	Ÿr.À	7,24	215		-		1
		• .	205					7.7.			-	-14	·.`				
		• • •	41			TGG	100	መልጥ	GTC	TCA	GCA	GTC	CAC	ACA	GAA	1:	325
ATA	GTC	AAG	CGT	TAC	AAC	TCG	ACC	TRI	Val	Ser	Ala	Val	His	Thr	Glu		•
Ile	Val	Lys	Arg	Tyr	Asn	Trp	225	-1-				230	i				*
* . * . *		220	60.			4								• •			4
				63.6	3 CIT	GGA	ልጥ ር	GAT	GCT	TTC	AAA	GAA	CTG	GCT	GCC	1.	373
GGG	AAT	TAC	GGC	CAG	Ser	GGA	Met	ASD	Ala	Phe	Lys	Glu	Leu	Ala	Ala		37
Gly	Asn	TYE	GTA	GIU	Ser	240	•••		•		245			• • •		,	-
	235	,				-	•				*		· 	 3 3 50	COM	- -	421
	C3.3	ccc	, CTPC	TGC	ATC	GCA	CAC	TCG	GAC	AAA	ATC	TAC	AGC	AAT	7 3 a	•	
CAG	Clu	GGC	Teu	CVS	Ile	GCA Ala	His	Ser	Asp	Lys	IIe	TAL	261	ASII	265		
250		<u></u>		-4-	255		* .			260				6			
230					ī						OTIC	ccc	CAG	CGG	CIT	1	469
GGC	GAG	AAG	AGC	TTT	GAC	CGG	CTC	CIG	CGT	AAA	Ten	Aro	Glu	Ara	Leu		
GIV	Glu	Lys	Ser	Phe	Asp	Arg	Leu	Leu	AFG	nys		****	,	280	Leu		**
			•	270							1		4.		•		
			•		· .	GTC		mmc	TCC	GAG	GGC	ATG	ACA	GTG	CGG	1	517
CCC	AAG	GCC	AGG	GTT	GTG	Val	TGC	Dho	LAG	Glu	Glv	Met	Thr	Val	Arg	145 -4	
Pro	Lys	Ala	Arg	Val	Val	ATT	Cys	290	, cys				295				
			285	, ,		. 1		3.0	. *			s					
						CGC	CGC	CTG	GGC	GTC	GTG	GGC	GAG	TTC	TCA Ser	1	.565
GGC	TTA	Cic	AGI	GCC	Mot	Ara	Ara	Leu	Gly	, Val	Val	Gly	Glu	Phe	Ser	• • • • • • • • • • • • • • • • • • • •	*
Gly	Lev	Let	ı ser	Ald	Mec	ura.	305					310)	***			
* 1		300	*	1 1			• 7 -						· · ·		ccc	· .	613
			N 2 CT	CAT	GGA	TGG	GCA	GAC	AGA	A GAT	' GAA	GTC	ATC	GAA	GGC	, -	*
CIC	All	GG/	, Sei	· Ast	Glv	Trp	Ala	Asp	Arc	Asp	Glu	Val	TIE	GTu	Gly	1 -	0
Ten	315	: GI				320					325				٠ - ڏ		
		0					• • •				336		CAG	тст	CCA		1661
ייבייי	GAG	GTY	GA	A GCC	AAC	: GGA	GGG	ATC	; AC	A ATA	AAG	. Tai	Gln	Ser	CCA Pro		
ጥህን	Gli	ı Va	l Gli	1 Ala	ASI	Gly	Gly	Ile	Ini	. 176	, Fris	, Dec			Pro 345		1 4
330	3	· · · ·	1.		222		1	F 4 3	1 F 2 1								
			:	•				. mene		- AAG	CTG	ÁGC	CTG	GAC	ACC		1709
GAG	GT	CAG	g TC	J. TTI	GAT	GAC	TAC	Dhe	LIC	i T.V.S	Leu	Arc	Leu	AS	Thr	*	*
Gli	ı Va	L Ar	g Se	C Pue	s waf) Asp	TYL	FIIG	35	5			; "	360)		
				JOU	<i>)</i>		·	**		*	0			•			1757
* 17	•		*		D (17/2/	- mm		GAG	TT	C: TG	G CA	CA!	r cgć	TT	CAG Gln	8	1757
AAC	CAC	A AG	G AA	T. CC.	r Han	Dhe	Pro	Gli) Ph	e Tr	o Gli	n His	Arc	r Phe	e Gln		
Ası	n Th	r Ar	g As	n Pro	,			370	- y	-	* .		375	5 .	G .		
* *			36					• • • •		*		<u> </u>			N GMC	· , · .	1805
*	. ~~	~ ~~	2	ጥ (GG)	A CA	CTC	TTC	GA	A AA	C CC	C AA	C TT	I AAC	s AAA	A GTG S Val	•	
TG.	T. C.G.	~ TA	71 DY	o Gl	y His	s Lev	i Let	1-G11	ı As	n Pro	O ASI	n Pho	B TA	a ny	s Val	•	·نــ بــر بــــ
· · cy	S AL	9 Le	0	,			385	5				39	IJ	. 0	٠.		
				•							m		C (2)	CAG	C AAA		1853
m~	C 3C	A GG	AAA	T GA	A AG	CTIC	GAI	A GA	A AA	C TA	T GT	1 - 1	n bei	o Se	C AAA r Lys		
7.6	e mh	r G1	v As	n Gl	u Se	r Lei	ı Glı	ı Gl	u As	n TY	. V₫.	K T GT:	" VO	F . J J	r Lys		
	. 39	5				70	_		•	-		•		•			
		<u> </u>		•		_ = -	'-		m ~~	ነ አጥ	6 60	A CA	T GG	G CT	G CAG	;	1901
AT	G GG	A TI	T GI	C AT	C AA	T GC	C AT	C. TA	1 66	,				•	G CAG		011557
		• •			•		,	*		•				S	UBST	TUTE	SHEET
								•			*	•	•				

				•									,	·		_		
	Mat~	Glv	Phe	Val	Ile	Asn	Ala	Ile	Tyr	Ala	Met	Ala	His	Gly	Leu	Gln		
	410	GLY	1110	7		415			, -		420	•			•	425		
				- ,	,				•	٠.				0 0			_	
	220	ATY	CAC	CAT	GCT	CTG	TGT.	CCC'	GGC	CAT	GTG	GGC	CTG	TGT	GAT	GCT	1	949
	Acn	Mot	His	His	Ala	Leu	Cys	Pro	Gly	His	Val	Gly	Leu	Cys	Asp	Ala	•	s.,
•	no			i	430		. •			435			•	•	440	8 .		,
		•		:														
	3 TG	222	COC	ATTY	GAT	GGC	AGG	AAG	CTC	CTG	GAT	TTC	CTC	ATC	AAA	TCC	, 1	.997
	Mat	Tue	Pro	Tle	ASD	Glv	Arg	Lys	Leu	Leu	Asp	Phe	Leu	Ile	Lys	Ser	,	• •
	Mec	ح ري.		445	*•÷₽			•	450	•			• .	455		•		
						•	•			`	8			0				
	mom.	ملحلمان	CTC	GGA	GTG	TCT	GGA	GAG	GAG	GTG	TGG	TTC	GAT	GAG	AAG	GGG	2	045
	ici	Dha	Unl	Glv	Val	Ser	Glv	Glu	Glu	Val	Trp	Phe	Asp	Glu	Lys	Glý		
	Ser	LINE	460				1	465					470				•	·:
							• •	4.7	•			*			· -			•
	C) M	COM		GGA	»GG	TAT	GAC	ATT	ATG	AAT	CTG	CAG	TAC	ACA	GAA	GCT Ala		2093
	GAT	GCT	CCC	Gly	Ard	Tur	Acn	Tle	Met	ASD	Leu	Gln	Tyr	Thr	Glu	Ala	. 0	
			PĻŪ	GTA	nr 9	-1-	480		1		-	485	- 0	•				
		۸75	• •						٠.		•		*		~	÷	•	
	330	CGC	ጥልጥ	GAC	TAT	GTC	CAC	GTG	GGG	ACC	TGG	CAT	GAA	GGA	GTG	CTG		2141
	yes.	7 20	Mar-	Aen	TVT	Val	His	Val	Glv	Thr	Trp	His	Glu	Gly	Val	Leu	٠,.	
	490		TAT	voh	1 -	495					500					505	-2	
											•	•						
		3 (79(1)	CATT	CAT	TAC	ΔΔΔ	ATC	CAG	ATG	AAC	AAA	AGC	GGA	ATG	GTA	CGA		2189
	AAT	ATT	GAT	Ben	TAC	Tue	Tle	Gln	Met	Asn	Lvs	Ser	Gly	Met	Val	Arg	٠. ا	* .
	ASI	TTE	. Asp	wab	510	Lys	440			515	. – 2	·	· · · · · -		520		** · · · ·	<u>.</u> ,
				•				• .		-							· .*	÷
•		cmc	mcc	, λ GΨ	GAG	CCT	TGC	איזייד	AAG	GGT	CAG	ATT	AAG	GTC	ATA	.CGG		2237
	TCT	GIG	760	VGI	Clu	Dro	CVE	T.A11	Tys	Glv	Gln	Ile	Lys	Val	Ile	Arg		
٠	Ser	AST	Cys	261	GIU	PLU	cys	200	530	,		٠	•	535				
		•		525	•							•						
,1			63.3	CMC	3.00	TOC	TGC	TICG	ATC	TGC	ACG	GCC	TGC	AAA	GAG	AAT	· · ·	2285
	AAA	GGA	GAA	. GTG	AGC	Lac	Cic	100	Tle	CVS	Thr	Ala	CVS	Lys	Glu	Asn		•
	Lys	GTĀ	GIU	AT	Ser	Cys	Cys	545	110				550	•		•	•	
			540			المراجعين		343	•,			. •				•		.:
						CAG	THE C	ACC	TGC	AGA	GCC	TGT	GAC	CTG	GGG	TGG		2333
	GAG	TIT	GIG	CAG	GAC	Clu	Dho	The		Ara	Ala	Cvs	ASD	Leu	Gly	Trp		
•	Glu			GIN	Asp	GIU	560	1117	- Cy 2	ALG		565			-			
	•	555		!			260				7.					• • •		
				;		-	3.03	ccc	men	GÀG	CCC	ATT	CCI	GTC	CGI	TAT		2381
	TGG	CCC	: AAC	GCA	GAG	CIC	ACA	- GGC	191	Clu	Dro	Tle	Pro	Val	Aro	TYF		•
	Trp	Pro	Asn	ALA	- Gin	Leu	Thi	GTA	Cys	GIU	580					Tyr 585	Ý	· /
	570	•		ij		575				7 0	٥٥٥		•			, -		* *
									3.000	A MT A	GCC	አ ጥር	GCC	TTT	TCI	TGC		2429
	CIT	GAG	TGG	AGI	GAC	ATA	GAA	TCI	ATC	, Ala	31=	Tle	Ala	Phe	Ser	TGC	- '	
	Leu	Glu	TIF) Ser	Asp	Ile	GIU	Ser	TTE	TTE	NTG	176	, nau		600	Cys		
					590)	-			595	•							
			•									3.000	י מינוי ר	·	CTC	TAC		2477
	CTG	GGC	ATC	CIC	GIG	ACG	CTG	Lalai	GIC	ACC	Cru	TIC	Dhe	Val	T.AL	TAC	•	
	Leu	Gly	II	Lev	. Val	. Thi	Leu	Phe	y val	Ini	. Ten	TTE	Pile	615		Tyr		
-	•	*		605	;-		•		610). ·. :			** *		- *			• • • • • • • • • • • • • • • • • • • •
	0				•							(A)	CORP.	י ייינור	ጥልባ	OTA-1		2525
	CGG	GAC	: AC	CCC	GTO	GTC	: AAA	TCC	TCC	AGI	AGG	GAU		, ,,,,,	ያ መደው የ	ATC Ile		
	Ara	Ast	Thi	Pro	Val	L Val	. Lys	Ser	. Sei	: 5 1	Arg	GIL		1-	- 1	r Ile	•	
	3	E	620					625	5				630	,				
				•		•	.:			. · · <u></u> .			n mm/	2 200	, CHI)	- ልጥሮ		2573
	ATT	CTC	GC	r GG1	AT	r TTC	CIC	GGC	TA	r GT	TGC	: CC	Dr.	, Mb-	· Tan	TILE		
	Ila	La	Ala	Gly	, Ile	a Ph	Let	ı Gly	Ty	r Val	r cha	FI	. PII	a TIII		ı Ile		
			1						•	. '		. ,		0	•	•		-

	•	635	•			ė	640		÷			645					
			CCT Pro	ACT Thr	THE	ACA Thr 655	TCC Ser	TGC Cys	TAC Tyr	CTC	CAG Gln 660	CGC Arg	CTC Leu	CTA Leu	GTT Val	GGC Gly 665	2621
-		TCT Ser	TCT Ser	GCC Ala	ATG Met 670	TGC Cys	TAC Tyr	TCT Ser	GCT Ala	TTA Leu 675	GTG Val	ACC Thr	AAA Lys	ACC Thr	AAT Asn 680	CGT Arg	2669
	ATT Ile	GCA Ala	CGC Arg	ATC Ile 685	CTG Leu	GCT Ala	GGC Gly	AGC Set	AAG Lys 690	AAG Lys	AAG Lys	ATC Ile	TGC Cys	ACC Thr 695	CGG Arg	AAG Lys	2717
	CCC	AGA Arg	TTC Phe 700	ATG Met	AGC Ser	GCT Ala	TGG Trp	GCC Ala 705	CAA Gln	GTG Val	ATC Ile	ATA Ile	GCC Ala 710	TCC	ATT Ile	CTG Leu	2765
	ATT Ile	AGT Ser 715		CAG Gln	CTA Leu	ACA Thr	CTA Leu 720	GTG Val	GTG Val	ACC Thr	TTG Leu	ATC Ile 725	ATC Ile	ATG Met	GAG Glu	CCT Pro	2813
	CCC Pro		CCC Pro	ATT	TTG	TCC Ser 735	TAC Tyr	CCG	AGT Ser	ATC Ile	AAG Lys 740	GAA Glu	GTC Val	TAC Tyr	CTT	ATC Ile 745	2861
	. ,	aat Asn	ACC Thr	AGC Ser	AAC Asn 750	CTG Leu	GGT Gly	GTA Val	GTG Val	GCC Ala 755		GTG Val	GGT Gly	TAC	AAT Asn 760	GGA Gly	2909
	CTC Leu	CTC Leu	ATC Ile	ATG Met 765	Ser	TGT	ACC Thr	TAC	TAT TYP 770	***	TTC	AAG Lys	ACC Thr	CGC Arg 775	AAC Asn	GTG Val	2957
	CCG Pro	GCC Ala	AAC Asn 780	TTC Phe		GAG Glu	GCT Ala	AAA Lys 785	-1-	ATC	GCC	TTC	ACC Thr 790	ATG Met	TAC	ACT	3005
	Thr	TGC Cys	ATC Ile	•	TGG Trp	CTG Leu	GCT Ala	LITE	GTT Val	CCC Pro	ATT	TAC Tyr 805		GCG	AGC Ser	AAC Asn	3053
٠	TAC Tyr	: AAG		: ATC	ACT Thr	ACC Thr 815	Cys	TTC Phe	GCG Ala	GTG Val	Ser 820		AGT Ser	GTG Val	ACG Thr	GTG Val 825	3101
	GCC Ala		GGG Gly	TGC Cys	: ATG : Met 830	Pne	ACI Thi	CCC Pro	AAG Lys	ATC Met	4 -	ATC	ATC	ATT	GCC Ala 840	Lys	3149
	CCI Pro	GAG	AGO Arg	J AAC J ASI 845	GTC		AG1	GCC Ala	TTC 2 Pho 850		ACC Thi	TCI	CAD ISA	GTI Val 855	GTC Val	CGC Arg	3197
	ATO Met	CAC His	GT(Va) 86(c GG1 L Gl3		GGC Gly	: AAI / Ly:	A CTO	G CC	- m <i>ci</i>	c cg(C TCC g Sei	C AAC r Asi 870	ACC Thi	Pho	C CTC	3245

AAC ATT TTC CGG AGA AAG AAG CCC GGG GCA GGG AAT GCC AAG AAG AGG AAC ATT TTC CGG AGA AAG AAG CCC GGG GCA GGG AAT GCC AAG AAG AGG AAC ATT TTC CGG AGA AAG AAG AAG AAG AAG AAG AAG AA	3293
CAG CCA GAA TTC TCG CCC AGC AGC CAG TGT CCG TCG GCA CAT GCG CAG Gln Pro Glu Phe Ser Pro Ser Ser Gln Cys Pro Ser Ala His Ala Gln 890 900	3341
CTT TGAAAACCCC CACACTGCAG TGAATGTTTC TAACGGCAAG TCTGTGTCAT	3394
Leu	•
GGTCTGAACC AGGTGGAAGA CAGGCGCCCA AGGGACAGCA CGTGTGGCAG CGCCTCTCTG	3454
TGCACGTGAA GACCAACGAG ACGGCCTGTA ACCAAACAGC CGTAATCAAA CCCCTCACTA	3514
AAAGTTACCA AGGCTCTGGC AAGAGCCTGA CCTTTTCAGA TGCCAGCACC AAGACCCTTT	3574
ACAATGTGGA AGAAGAGGAC AATACCCCTT CTGCTCACTT CAGCCCTCCC AGCAGCCCTT	3634
CTATGGTGGT GCACCGACGC GGGCCACCCG TGGCCACCAC ACCACCTCTG CCACCCCATC	3694
TGACCGCAGA AGAGACCCCC CTGTTCCTGG CTGATTCCGT CATCCCCAAG GGCTTGCCTC	3754
TGACCGCAGA AGAGACCCCC CIGITOGICO CTCCTCTCCC GCAGCAGCAG CCACAGCAGC CGCCCCCTCA GCAGCCCCCG CAGCAGCCCA	3814
AGTCCTGAT GGACCAGCTG CAAGGCGTAG TCACCAACTT CGGTTCGGGG ATTCCAGATT	3874
TCCATGCGGT GCTGGCAGGC CCGGGGACAC CAGGAAACAG CCTGCGCTCT CTGTACCCGC	3934
TCCATGCGGT GCTGGCAGGC CCGGGGACAC CCCGCCTCCA CCTGAGCACC TTCCAGGAGG	3994
CCCCGCCTCC GCCGCAACAC CTGCAGATGC TGCCCCTGCA CCTGAGCACC TTCCAGGAGG	4054
AGTCCATCTC CCCTCCTGGG GAGGACATCG ATGATGACAG TGAGAGATTC AAGCTCCTGC	4114
AGGAGTTCGT GTACGAGCGC GAAGGGAACA CCGAAGAAGA TGAATTGGAA GAGGAGGAGG	4174
ACCTGCCCAC AGCCAGCAAG CTGACCCCTG AGGATTCTCC TGCCCTGACG CCTCCTTCTC	
CTTTCCGAGA TTCCGTGGCC TCTGGCAGCT CAGTGCCCAG TTCCCCCGTA TCTGAGTCGG	4234
TCCTCTGCAC CCCTCGAAAT GTAACCTACG CCTCTGTCAT TCTGAGGGAC TACAAGCAAA	4294
GCTCTTCCAC CCTGTAGTGT GTGTGTGTT GTGGGGGGGGGG	4354
CCAGAGATGC CAAGGAGTGT CAACCCTTCC AGAAATGTGT AGAAAGCAGG GTGAGGGATG	4414
CCAGAGATGC CAAGGAGTGT CHACCOURTS AAAAAAAAAAA TGCTGCGGCT GCCTTAAAGA GGGATGGAGG ACCACGGTCT GCAGGGAAGA AAAAAAAAAA	4474
GGGATGGAGG ACCACGGTCT GCAGGGAAGA AAMAGATTAT	4534
AGGAGAGGA CGATGCCAAC TGAACAGTGG TCCTGGCCAG GATTGTGACT CTTGAATTAT	4594
TCAAAAACCT TCTCTAGAAA GAAAGGGAAT TATGACAAAG CACAATTCCA TATGGTATGT	4654
AACTTTTATC GAAAAAATA ATAAAACGTA AAAATAAAAT	
TTTGCTCAAT CGTGCATACA TATATCTGCC CACACTCCCG TGGTAAAACT AGAAGCGAAG	471
CAGGCCCTGC GATGGTGCCA ACTGAATCCT AAGTTCATCA TCCTAGTGAG CAGATGGAGA	477
SUBSTITUTE SHEET	

	4024
GASGCAGGA GGCGAGAGGG CAGGAGGCGG GGGTAGG	TTC GGACAACAGC TCCCATCTCA 4834
GACCTTGACT GTGCTGAGTC TTCAGACTCC TGGACTA	AGG AAGACCCGGG GACTGACCTT 4894
ATGAGGGTCC CTTTCCACTG CTGTGATCCA TTGCCAG	CCT GTAGTCACCC GGGATAAAGG 4954
CACAGTAACC TTTTGCATTC CTGTGATTCC CTGTGTT	TAA GGAAAAGGAA AGTATGAGCA 5014
AAGCTATCAC CAAAAAGAGC GCCATTAGAA GTTACGG	GGG AGAAAAAAG AGAAGCAAGA 5074
TGATATATAA GCACAGGGCC TTGAACAAGG TGAGCGT	GCT TCACAGATTC CGTATTAATG 5134
TACAGATACT TTTGGAGAGG AGAAAGATAA CAAGGAG	STGT CAGGCCGTTT GTGAACTCAC 5194
TTGCACTGTG CCAACCAGGT TCTCCGCTGC CCTTCAC	
(2) INFORMATION FOR SEQ ID NO:17:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 906 amino acid	
(B) TYPE: amino acid (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: protein	
(xi) SEQUENCE DESCRIPTION: SEQ ID	NO:17:
Met Val Arg Leu Leu Leu Ile Phe Pro	Met Ile Phe Leu Glu Met 15
1	
Ser Ile Leu Pro Arg Met Pro Asp Arg Lys	•
Ser Ser Gln Arg Ser Val Ala Arg Met Asp	Gly Asp Val Ile Ile Gly
35	
Ala Leu Phe Ser Val His His Gln Pro Pro	Ala Glu Lys Val Pro Glu 60
Arg Lys Cys Gly Glu Ile Arg Glu Gln Tyr	Gly Ile Gln Arg Val Glu
65	
Ala Met Phe His Thr Leu Asp Lys Ile Ass	Ala Asp Pro Val Leu Leu 95
85.	
Pro Asn Ile Thr Leu Gly Ser Glu Ile Ard	g Asp Ser Cys 119 mas 555
Ser Val Ala Leu Glu Gln Ser Ile Glu Pho	e Ile Arg Asp Ser Leu Ile
115	ş •
Ser Ile Arg Asp Glu Lys Asp Gly Leu Ass	n Arg Cys Leu Pro Asp Gly 140
130	
Gln Thr Leu Pro Pro Gly Arg Thr Lys Ly	155

Gly	-Pro	Gly	Ser	Ser 165	Ser	Val	Ala	Ile	Gln 170	Val	Gln	Asn	Leu	Leu 175	Gln.
		•	180				•	105	•		Thr				•
		105					200		•						Asp
	210		*		ė	215	•				Lys 220				11.5
225	·			•	230					235	Tyr	•	-		
		•		245		*			250		Gly				,
	•		260	. :				265			:				Arg
	:	275				`	280			•	Ala			•	
	290			•		295			•		Leu 300				
30	5				310		•			315					Trp 320
	•		•	325	•				330					•	
			340	0				345							Asp_
•		355	5				300				•	0			Phe
	37	0	-	.		375					300	•	•		Leu
38	5		•		390						•			_	400
				40:	•			,	47/	•			•		
			420)			• =	42:)						ı Cys
		43	5;	• .			. 441	,	· ·			7.7.	7.		y Arg
L	's Le 45		u As	p Pho	e Leu	1 Île 455	Lys	s Sei	r Se:	r Pho	e Val	l Gly	y Va	l Se	r Gly

465			•		470			•		-	Pro				
			7	485					•••		Tyr				
		•	500		111		1	200			Asp	٠.	0		
		515	;				224				Cys	•		•	
•	530					232									Cys
545			•		220			•				•			
				565	* -	•					Asn			•	
			580					500			Trp			•	
		595	Ŧ				000		*		Ile				
	610)				010					•				Lys
625	,		1		930	,									
				543	*				, T		Pro		•		•
	4		660)											Tyr
	•	6.75	; ;	v.			000	'	•					•	Gly
	691)	. :			03.									Trp
Ala 70!	a Gla	n Val	l II	e Ile	2 Ala 710	s Sez	r Ile	Leu	ı Ile	71!	r Val				720
				72:	5				•		t Pr	-	40	735	•
Pr	o Se	r Il	e Ly:	0				, ••							gly
Va	1 Va	1 Ala 75		۷a	1 G1	у Ту	r Asi 760	n Gly	y Le	u Le	u Il	76	t Sei 5	r Cyt	3 Thr

Tyr Tyr Ala Phe Lys Thr Arg Asn Val Pro Ala Asn Phe Asn Glu Ala 770 775 780

• 10		•			, .						_		-			•	
785			•		790				. ·	Cys 795					2		
Phe	Val	Pro	Ile	Tyr 805	Phe	Gly	Ser	Asn	Tyr 810	Lys	Ile	Ile	Thr	Thr 815	Cys		
Phe	Ala	Val	Ser 820	Leu	Ser	Val	Thr	Val 825	Ala	Leu	Gly	Cys	Met 830	Phe	Thr	•	•
Pro	Lys	Met 835		Ile	Ile	Ile	Ala 840	Lys	Pro	Glu	Arg	Asn 845	Val	Arg	Ser		
Ala	Phe 850		Thr	Ser	Asp	Val 855	Val	Arg	Met	His	Val 860	Gly	Asp	Gly	Lys		
Leu 865		Cys	Arg	Ser	Asn 870	Thr	Phe	Leu	Asn	Ile 875	Phe	Arg	Arg	Lys	Lys 880		•
		Ala	Gly	Asn 885	Ala	Lys	Lys	Arg	Gln 890	Pro	Glu	Phe	Ser	Pro 895	Ser	· •	
Ser	Gln	Cys	Pro 900	Ser	Ala	His	Ala	Gln 905	Leu	· · ·	•		-		•		
(2)	(2) INFORMATION FOR SEQ ID NO:18: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 4095 base pairs																
															9		
(ii) MOLECULE TYPE: CDNA (vii) IMMEDIATE SOURCE: (B) CLONE: SN30																	
(ix) FEATURE: (A) NAME/KEY: CDS (B) LOCATION: 4633198															* * * * * * * * * * * * * * * * * * *		
	•									NO: 18					,	•	
															rcccc		60
															AGTGA		120
															CCCTT		180
AC	AGCA(GAC	ACA	GAAA'	rct (GCC.	PTCA(GT A	CTTT	GGGA)	A.AA	ggat(CTGA	GAC	CTCCT	GG ma	30
AG	CTCI	GACC	ACT	GGCTY	STC 3	ATCT	GTGG	CT C	TGGC	CIGI	g TG	GGCC	ACTG	AGC	TCTAC	TU V	3 00
	•					CII	B \$7	riT!	ITE	SH	EET	Γ .			-	-	

92	
AALCATTAAA GAGGAGGAGG GGAGATCTGT GGAATGGGCC ACCCCGTTGG CCTGCTGC	CAT 360
AAACATTAAA GAGGAGGAGG CONGINE OO TACTGAACCT GCGGGCCCA CACGTGCCCA GATCATGGGA CCCAGGGCCT GCTAGGGC	CTA 420
GGAGCGGGGC CCAGTATTCA TGGGTCTCTA GGCCTTTCCG AA ATG TCC GGG AAG GGAGCGGGGC CCAGTATTCA TGGGTCTCTA GGCCTTTCCG AA ATG TCC GGG AAG Met Ser Gly Lys	474
	* "
GGA GGC TGG GCC TGG TGG GCC CGG CTG CCC CTC TGC CTA CTC CT Gly Gly Trp Ala Trp Trp Ala Arg Leu Pro Leu Cys Leu Leu Le 10	C 522 u 0
AGC CTT TAT GCC CCC TGG GTG CCT TCA TCC TTG GGA AAG CCC AAG GG Ser Leu Tyr Ala Pro Trp Val Pro Ser Ser Leu Gly Lys Pro Lys Gl	T 570 Y
CAC CCC CAC ATG AAC TCT ATC CGA ATT GAC GGG GAC ATC ACA CTG GG His Tro His Met Asn Ser Ile Arg Ile Asp Gly Asp Ile Thr Leu Gl 45	A 618 Y
GGC CTG TTT CCC GTC CAC GGC CGT GGC TCT GAG GGT AAG GCC TGC GC Gly Leu Phe Pro Val His Gly Arg Gly Ser Glu Gly Lys Ala Cys Gl 65	
GAG CTG AAG AAG GAG AAA GGC ATC CAC CGC CTG GAG GCC ATG CTG T Glu Leu Lys Lys Glu Lys Gly Ile His Arg Leu Glu Ala Met Leu Pl 75	
GCC CTG GAC CGC ATC AAC AAT GAC CCG GAC CTA CTG CCC AAC ATC AAA Leu Asp Arg Ile Asn Asn Asp Pro Asp Leu Pro Asn Ile T	
TTG GGC GCC CGC ATT CTG GAC ACC TGC TCG AGG GAC ACC CAC GCC C Leu Gly Ala Arg Ile Leu Asp Thr Cys Ser Arg Asp Thr His Ala L 105	
GAG CAG TCA CTG ACC TTT GTG CGG GCG CTC ATC GAG AAG GAC GGC A Glu Gln Ser Leu Thr Phe Val Arg Ala Leu Ile Glu Lys Asp Gly T	
GAG GTC CGC TGG GGC AGG CGG GGC CCG CCC ATC ATC ACC AAG CCC GGU Val Arg Cys Gly Arg Arg Gly Pro Pro Ile Ile Thr Lys Pro Glu Val Arg Cys Gly Arg Arg Hand 145	•
CGA GTG GTG GGT GTC ATT GGA GCT TCG GGG AGC TCC GTC TCG ATC ATG Val Val Gly Val Ile Gly Ala Ser Gly Ser Ser Val Ser Ile 1	•
GTG GCC AAC ATC CTC CGC CTC TTC AAG ATC CCT CAG ATC AGC TAT (Val Ala Asn Ile Leu Arg Leu Phe Lys Ile Pro Gln Ile Ser Tyr 170 175	
TCC ACG GCC CCT GAC TTG AGT GAC AAC AGC CGC TAT GAC TTC TTC Ser Thr Ala Pro Asp Leu Ser Asp Asn Ser Arg Tyr Asp Phe Phe 185	
CGG GTG GTG CCC TCA GAC ACA TAC CAG GCC CAG GCC ATG GTG GAT	ATT 1098
- CUFFT	•

STRETTUTE SHEET

	•																•	
•			Pro 200			. •		203			٠.							
Va.	l Arg	Ala 215	CTC Leu	Lys	Trp	ASN	220	Val	261	****		225		, 	•	*	1146	
AG Se	C TAC r Tyr 230	Gly	GAG Glu	AGT Ser	GGT Gly	GTG Val 235	GAG Glu	GCC Ala	TTT	ATC Ile	CAG Gln 240	AAG Lys	TCC Ser	CGÀ Arg	GAG Glu	*	1194	
AA As 24	n Gly	GGT Gly	GTG Val	TGC Cys	ATT Ile 250	GCC Ala	CAG Gln	TCG Ser	GTG Val	AAG Lys 255	ATT Ile	CCA Pro	CGG Arg	GAA Glu	CCC Pro 260		1242	
AA Ly	G ACC	GGG Gly	GAG Glu	TTC Phe 265	GAC Asp	AAG Lys	ATC	ATC Ile	AAA Lys 270	CGC	CTA Leu	CTG Leu	GAA Glu	ACA Thr 275	TCC		1290	
	T GC	C AGG A Arg	GGT Gly 280	Ile	ATC Ile	ATC Ile	TTT Phe	GCC Ala 285	ASII	GAG Glu	GAT Asp	GAC	ATC Ile 290	AGG	AGG Arg	œ.	1338	:
GI Va	G TT	G GAG u Glu 295	GCA Ala	GCT	CGC	AGG Arg	GCC Ala 300	ASI	CAG Gln	ACC Thr	GGC Gly	CAC His 305	TTC Phe	TTT	TGG Trp	8.	1386	
AT Me	G GG t Gl	y Sei	GAT Asp	AGC	TGG	GGC Gly 315	Ser	AAG Lys	AGT Ser	GCC Ala	CCT Pro 320	GTG Val	CTG Leu	CGC	CTT		1434	
G.	AG GA Lu Gl 25	G GT(u Va	GCC L Ala	GAG Glu	GGC Gly 330	ATG	GTC Val	ACC	ATT : Ile	CTC Leu 335	CCC	AAG Lys	AGG Arg	ATG Met	TCT Ser 340		. 1482	
		A GGG	G TTC y Phe	GAC Asp 345	Arg	TAC	TTC Phe	TCC Ser	AGC Ser 350	ur à	ACG Thr	CTG Leu	GAC Asp	AAC Asn 355			1530	
A.	sg co	ic AA	C ATO	a Tri	TTI Phe	GCC Ala	GAG Glu	TTC Phe 365	Trp	GAG Glu	GAC Asp	AAC Asr	TTC Phe 370		TGC		1578	
A L	AG TI YS Le	G AG au Se 37	r Arg	CAC His	C GCG	CTC	AAG Lys	, Th	G GGA S Gly	AGC Ser	CAC His	385		AAG Lys	TGC Cys		1626	
A T	hr As	AC CG on Ar	A GAO	G CGG	C ATO	GGG Gly 39!	A GTI	GA(C TCG p Ser	GCC Ala	TAT TY1		CAG Glr	GAC	GGG Gly	*	1674	
L			G TT n Ph	C GT B Va	G AT:	e asi	C GCT p Ala	r GT a Va	G TAC	GCC Ala		G GGG	C CAC y His	C GCC	CTG Leu 420	0	1722	
		CC AT	G CA t Hi	C CG	T GA	C CTO	G TG	r cc s Pr	O GI	C CGC	_	A GG	À CT(y Lei	C TGO	C CCT s Pr		1770	

				*						•	1	3						
					425		.*		. , '	430		•			435		.0	
	CGC	ATG Met	Asp	CCC Pro	GTG Val	GAT (Asp (GGC A		CAG	CTG Leu	CTT Leu	AAG Lys	TAC Tyr	ATC Ile 450	AGG Arg	AAC Asn	` <u></u> 1	1818
•	GTC Val	AAC Asn			GGC Gly	ATT Ile	UT A	GGG Gly 460	AAC Asn	CCT Pro	GTA Val	ACC Thr	TTC Phe 465	AAT Asn	GAG Glu	AAC Asn		1866
*	GGA Gly	GAC Asp 470	•	CCG Pro	GGG Gly	CGC Arg	TAC Tyr 475	GAC Asp	ATC Ile	TAC Tyr	CAG Gln	TAC Tyr 480	CAA Gln	CTG Leu	CGC	AAT Asn		1914
•	Gly 485	TCG Ser	Ala	Gļu,	TYI	190	VOI		U -1		495	*	,		•	•	0	1962
	Leu	Arg	Ile	GIU	CGG Arg 505	Mer				510	:				515			2010
	Arg	Ser	Ile	520	AGT Ser	Ten	PIO	-0,0	525			. • .	. :::	530	-	· .	- - <u>-</u> -	
	Val	Lys	535 535	Met	WIG	Cyu		540					545		٠,.	TAC		2106
	Gln	Ty:	Gln	Val	Asp	ALG	555			_		560				ATG Met		
ė	Arg	Pro	Thr	GIU	ASI	570	1114				575	5				Lys 580		2202
	Lev	Glu	ı Tip	AST L	- 585	PLO	115			590					59	C GTG a Val		2250
	Va]	L Gl	y Ile	5 Ala 3 600	J	THE	Ten		605	5				610	L Are	C TAC		2298
	Ası	n Asi	p Thi 615	r Pro) IT€	. vai	, rye	620				- . ,	625	5		c GTG r Val	· .	2394
	Let	Le 63	u Ala O	g GI	À TT	FILE	635	5	3	-	: ,	64	0			G ATC		2442
	GC Al 64	a Gl	G CC	G GA	C CTO	650)			•	65	5	C ATO	C TT Ph	c CI	A GGG u Gly 660	···	6774
						S	UB	5 Tľ	TUT	E :)HE		•	•				•
				•		, —					,							

		, AAA	Les	s, GMC		•			*					SUE	STIT	TUTE	SHEET	
	<u>icc</u>			GAG	CI	TG1	r GAC	S AAC	CIC	G GAG	ACC	CCI	GCG	CTC	GCT	ACC	31	62
	Met	S 1 870	Ası	Lys	Phe	Thi	G11 87) TĀS	Gly	y Asr	Phe	880	PLU) Ast	ı Gly	r Glu	(4)	
-	3.000	, mee	855 		. Trees	· ACA	CAC	2 226	: GG0	C AAC	TTC	AGG	CCC	LAA :	GGG	GAA Glu	31	14
	GTG Val	Pro	Lys	Arg	Lys	Arg	Sei	Lev 860	r ras	Ala	Val	Val	Thr 865		Ala	Thr		
		*: 	*	840			• BCG	n (m)(A GCC	GTIC	GTC	ACC	GCC	GCC	ACC	30	66
	Leu	TAC	Met	Pro	Lys	Val	Tyr	Ile	Ile 845	Leu	Phe	His	Pro	Glu 850		Asn		
					825	CMC	• ጥልብ	• ልጥር	ልጥሮ	CTC	TTC	CAC	CCG	GAG	CAG	AAC	30:	18
	ACA Thr	CTG Leu	ACG Thr	GTÇ Val	Ser	· Val	Ser	Leu	Ser	: GCI : Ala : 830	2er	Val	Ser	Leu	Gly 835	ATG Met	10	st ²
	805	-,		. 1		810				, .	913	-	-				297	70
	Phe	Phe	GGC Gly	ACC Thr	TCA Ser	Gln	Ser	GCC	GAC	Lys	Leu	TYL	ATC Ile	Gln	Thr	Thr 820	232	
-		790		ì			795				• ,	800					292	22
	TTC Phe	ACC	ATG Met	TAC	ACC Thr	ACC Thr	cys	ITE	GTC Val	TGG Trp	CTG Leu	WIG	TTC Phe	ATC	Pro	Ile	287	7 4€
	٠.		775			•		780	.*			•	,05				207	, , ,
	AAG	ACC	CGA	GGC	GTG Val	CCC	GAG Glu	ACC Thr	TTC Phe	AAC	GAG Glu	GCC Ala	AAG Lys	CCC	ATC Ile	GGC Gly	282	6
	Leu	Gly	Tyr	5er 760	Met	Leu	Leu	Met	765	THE	Cys	THE	VAI	770				
	CTG	GGC	TAC	AGC	ATG	CTG	CTG	ATG	GTC	ACG	TGT	ACT	GTG Val	TAC	GCC Ala	ATC Ile	277	8
	Arg	Gly	Val	Leu	Lys 745	Cys	Asp	Ile	Ser	Asp 750	Leu	Ser	Leu	ITE	755			
•	100	GGC	GTG	CTC	226	TGC	GAC	ATC	TCG	GAC	CTG	TCC	CTC	ATC	TGC	CTG	273	o .
:	TCG Ser 725	GTG al	Val	Asp	Phe	Gln 730	Asp	Gln	Arg	Thr	Leu 735	Asp	Pro	Arg	Phe	Ala 740		
		710 ·	ama .	CAC	ww.	CAG.	GAC	CAA	CGG	ACA	CTT	GAC	CCC	CGC	TTT	GCC	268	2
-	CTG Leu	Gln	Leu	CTC	GGC	ATC Ile	TGC Cys 715	GTG Val	TGG	Phe	Val	Val	GAC Asp	Pro	Ser	His		,
			695					700		•		*				Single	263	4
	TTC Phe	ATC Ile	AGC Ser	CCG	GCC Ala	TCG Ser	CAG Gln	CTG Leu	GCC Ala	ATC Ile	ACC Thr	TTC Phe	ATC Ile	CTC Leu	ATC Ile	TCC Ser	258	6
	Ile	Tyr	Arg	Ile 680	Phe	Glu	GIN	GTA '	685	Arg	Ser	467		690			¥ .	
	ATT	TAC	CGC			GAG	CAG	GGC		CGG	TCG	GTC	AGT	GCC	CCG	CGT	2538	3
	CTC	GCC	ATG Met	Ser	Ile 665	Ser	Tyr	Ala	ALA	Leu 670	Leu	Thr	Lys	T ***	Asn 675	Arg		
	·		\		> mc	N G C	ጥልሮ	ccc	GCC	CTG	CTG	ACC	AAG	ACC	AAC	CGC	2490)

Ala-Lys Ser Glu Leu Cys Glu Asn Leu Glu Thr Pro Ala Leu Ala Thr 885 890 895	*
AAA CAG ACC TAC GTC ACC TAC ACC AAC CAT GCC ATC TAGCCGGGCC Lys Gln Thr Tyr Val Thr Tyr Thr Asn His Ala Ile 905	3208
GCGGAGCCAA-GCAGGCTAAG GAGCCACAAC CTCTGAGGAT GGCACATTGG GCCAGGGCCG	3268
TTCCCGAGGG CCCTGCCGAT GTCTGCCCGC CTCCCGGGCA TCCACGAATG TGGCTTGGTG	3328
CTGAGGACAG TAGAGACCCC GGCCATCACT GCTGGGCAAG CCGTGGTGGG CAACCAGAGG	3388
AGGCCGAGTG GCTGGGGCAG TTCCAGGTTA TGCCACACAC AGGTCTTCCT TCTGGACCAC	3448
TGTTGGCCCA GCCCCAAAGC ACAGGGGCTC GGTCTCCAGA GCCCAGCCCT GGCTTCCTCT	3508
TGTTGGCCCA GCCCCAAAGC ACHGCGCGGG TGACCCCGGT TGGTCCCTGC CCCGTCTTTA	3568
CGTTTCTCTT CCGTCTTTGC TCTGCATGTG TTGTCTGTTT GGGCCCTCTG CTTCCATATT	3628
THECATTET GETECTIGE TELECRICAL ATCTGCCCTG CCCCTGCCC CTCCTCCCTG	3688
AGCTGCCCCA TCCCCGCCAT CATTTCTCT TCTGTTCCCC CTCGATCTCA TTTCCTACCA	3748
GCCTTCCCCC TACTTGGCTT CATCCACCAA CTCTTTCACC ACGTTGCAAA AGAGAAAAAA	3808
AAAGGGGGG GGGAATCACC CCCTACAAAA AAGCCCAAAC AAAAACTAAT CTTGAGTGTG	3868
TTTCGAAGTG CTGCGTCCTC CTGGTGGCCT GTGTCTCCCT GTGCCTGCAG CCTGTCTGCC	3928
CGCCCTACCC GTCTGCCGTG TGTCCTGCCC CCCCCGCCTG CCCGCCTTGC CCTTCCTGCT	3988
AACGACACGG AGTTCAGTGC CTGGGTGTTT GGTGATGGTC TCTGATGTGT AGCATGTCTG	4048
TITITATACC GAGAACATTT CTAATAAGA TAAACACATG GTTTTGC	409

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 912 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:
- Met Ser Gly Lys Gly Gly Trp Ala Trp Trp Trp Ala Arg Leu Pr Leu
 1 10 -15
- Cys Leu Leu Ser Leu Tyr Ala Pro Trp Val Pro Ser Ser Leu Gly 25
- Lys Pr Lys Gly His Pro His Met Asn Ser Ile Arg Ile Asp Gly Asp
 45
 SUBSTITUTE SHEET

			•													
		∵ 50		Gly			22		*							
	65			Gly	÷	70	*		Ē		,, ,		٠.			
	Ala	Met	Leu	Phe	Ala 85	Leu	Asp	Arg	Ile	Asn 90	Asn	Asp	Pro	Asp	Leu 95	Leu
•	Pro	Asn	Ile	Thr	Leu	Gly	Ala	Arg	Ile 105	Leu	Asp	Thr	Cys	Ser 110	Arg	Asp
•	Thr	His	Ala 115	Leu	Glu	Gln	Ser	Leu 120	Thr	Phe	Val	Arg	Ala 125	Leu	Ile	Glu
	Lys	Asp		Thr	Glu	Val	Arg 135	Cys	Gly	Arg	Arg	Gly 140	Pro	Pro	Ile	Ile
	Thr 145	Lys	Pro	Glu	Arg	Val 150	Val	Gly	Val	Ile	Gly 155	Ala	Ser	Gly	Ser	Ser 160
	Val	Ser	Ile	Met	Val 165	Ala	Asn	Ile	Leu	Arg 170	Leu	Phe	Lys	Ile	Pro 175	Gln
	Ile	Ser	Tyr	Ala 180	Ser	Thr	Ala	Pro	Asp 185	Leu	Ser	Asp	Asn	Ser 190	Arg	Tyr
	Asp	Phe	Phe 195	ser	Arg	Val	Val	Pro 200	Ser	Asp	Thr	Tyr	Gln 205	Ala	Gln	Ala
	Met	Va]	Asp	Ile	Val	Arg	Ala 215	Leu	Lys	Trp	Asn	Tyr 220	. Val	Ser	Thr	Leu
•	Ala 225		: G1:	ı Gly	Ser	Tyr 230	Gly	, Glu	Ser	Gly	Val 235	Glu	Ala	Phe	Ile	G1n 240
	Lys	s Sei	r Arc	g Glu	Asr 245	Gly	Gly	y Val	. Cys	Ile 250	Ala	Gln	Ser	Val	. Lys 255	Ile
	Pro	o Ar	g Gli	1 Pro 260	Lys	Thr	Gly	y Glu	Phe 265	Asp	Lys	Ile	Ile	Lys 270	Arg	Leu
	Let	j Gli	u Th: 27	r Ser	Ası	n Ala	Arg	g Gly 280	, Ile	Ile	Ile	Phe	Ala 285	ASI	Glu	Asp
	Asj	o Il 29		g Arg	Val	l Lev	Gl: 29:	u Ala 5	Ala	Arg	, Arc	300	Ası	ı Glı	1 Thr	Gly
	Hi:	s Ph		e Trp	Me	t Gly 310	, Se:	r Asj	Ser	Tr	Gly 31:	y Sei 5	Ly	s Se	r Ala	320
			u Ar	g Lev	32:	u Glu	ı Va	l Ala	a Glu	1 Gly 330	/ Ala	a Va	l Th	r Il	e Let 339	Pro
	Ly	s Ar	g Me	t S 1	. Va	10.00	g Gl	y Ph	e Ası 34	o Arq	TY	r Ph	e Se	r S	r Arq	g Thr

	• 1			٠.										_			
			355			Arg		200			•			•			
	•	370				Leu	3/5	-						•			
	Ile 385	Lys	Lys	Cys	Thr	Asn 390	Arg	Glu	Arg	Ile	Gly 395	Gļn	Asp	Ser	Ala	Tyr 400	
	Glu	Gln	Glu	Gly	Lys 405	Val	Gln	Phe	Val	Ile 410	Asp	Ala	Val	Tyr	Ala 415	Met	
	Gly	His	Ala	Leu 420	His	Ala	Met	His	Arg 425	Asp	Leu	Cys	Pro	Gly 430	Arg	Val	(4)
•	Gly	Leu	Cys 435	Pro	Arg	Met	Asp	Pro 440	Val	Asp	Gly	Thr	Gln 445	Leu	Leu	Lys	
	Tyr	Ile 450	Arg	Asn	Val	Asn	Phe 455	Ser	Gly	Ile	Ala	Gly 460	Asn	Pro	Val	Thr	,
	Phe 465	Asn	G1u	Asn	Gly	Asp 470	Ala	Pro	Gly	Arg	Tyr 475	Asp	Ile	Tyr	Gln	Tyr 480	•
ī	Gln	Leu	Arg	Asn	Gly 485	Ser	Ala	Glu	Tyr	Lys 490	Val	Ile	Gly	Ser	Trp 495	Thr	
	Asp	His	Leu	His 500	Leu	Arg	Ile	Glu	Arg 505	Met	Gln	Trp	Pro	Gly 510	Ser	Gly	
	Gln	Gln	Leu 515	Pro	Arg	Ser	Ile	Cys 520	Ser	Leu	Pro	Cys	Gln 525	Pro	Gly	Glu	
	Arg	Lys 530	Lys	Thr	Val	Lys	Gly 535	Met	Ala	Cys	Cys	Trp 540	His	Cys	Glu	Pro	•
	Cys 545	Thr	Gly	Tyr	Gln	Tyr 550	Gln	Val	Asp	Arg	Tyr 555	Thr	Cys	Lys	Thr	Cys 560	
٠.		Tyr	Asp	Met	Ar g . 565	Pro	Thr	Glu	Asn	Arg 570	Thr	Ser	Cys	Gln	Pro 575	Ile	
	Pro	Ile	Val	Lys 580	Leu	Glu	Trp	Asp	Ser 585	Pro	Trp	Ala	Val	Leu 590	Pro	Leu	
	Phe	Leu	Ala 595	Val	Val	Gly	Ile	Ala 600	Ala	Thr	Leu	Phe	Val 605	Val	Val	Thr	
	Phe	Val 610	Arg	•	Asn	Asp	Thr 615	Pro	Ile	Val	Lys	Ala 620	Ser	Gly	Arg	Glu	
	Leu 625	sr		Val	Lev	Leu 630	Ala	Gly	Ile	Phe	Leu 635	Cys	Туг	Ala	Thr	Thr 640	
			Met	: Ile	Ala 645	Glu	Pro) Asp	Leu	Gly 650	Thr	Cys	Ser	Lev	Arg 655	Arg	
		•			,						•						

Ile Phe Leu Gly Leu Gly Met Ser Ile Ser Tyr Ala Ala Leu Leu Thr 660 Lys Thr Asn Arg Ile Tyr Arg Ile Phe Glu Gln Gly Lys Arg Ser Val 6**80** -Ser Ala Pro Arg Phe Ile Ser Pro Ala Ser Gln Leu Ala Ile Thr Phe Ile Leu Ile Ser Leu Gln Leu Leu Gly Ile Cys Val Trp Phe Val Val 710 705 Asp Pro Ser His Ser Val Val Asp Phe Gln Asp Gln Arg Thr Leu Asp .730 Pro Arg Phe Ala Arg Gly Val Leu Lys Cys Asp Ile Ser Asp Leu Ser Leu Ile Cys Leu Leu Gly Tyr Ser Met Leu Leu Met Val Thr Cys Thr Val Tyr Ala Ile Lys Thr Arg Gly Val Pro Glu Thr Phe Asn Glu Ala Lys Pro Ile Gly Phe Thr Met Tyr Thr Thr Cys Ile Val Trp Leu Ala 795 790 Phe Ile Pro Ile Phe Phe Gly Thr Ser Gln Ser Ala Asp Lys Leu Tyr 810 805 Ile Gln Thr Thr Leu Thr Val Ser Val Ser Leu Ser Ala Ser Val 825 Ser Leu Gly Met Leu Tyr Met Pro Lys Val Tyr Ile Ile Leu Phe His 835 Pro Glu Gln Asn Val Pro Lys Arg Lys Arg Ser Leu Lys Ala Val Val 855· 850 Thr Ala Ala Thr Met Ser Asn Lys Phe Thr Gln Lys Gly Asn Phe Arg 875 870 Pro Asn Gly Glu Ala Lys Ser Glu Leu Cys Glu Asn Leu Glu Thr Pro 885 Ala Leu Ala Thr Lys Gln Thr Tyr Val Thr Tyr Thr Asn His Ala Ile 905 900

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2426 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

_ (ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE: (B) CLONE: SR13

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:	
CCCAACATCA CGTTGGGCGC CCGCATTCTG GACACCTGCT CGAGGGACAC CCACGCCCTG	60
GAGCAGTCAC TGACCTTTGT GCGGGCGCTC ATCGAGAAGG ACGGCACGGA GGTCCGCTGC	120
GGCAGGCGGG GCCCGCCCAT CATCACCAAG CCCGAACGAG TGGTGGGTGT CATTGGAGCT	180
TCGGGGAGCT CCGTCTCGAT CATGGTGGCC AACATCCTCC GCCTCTCAA GATCCCTCAG	240
ATCACCTATG CCTCCACGGC CCCTGACTTG AGTGACAACA GCCGCTATGA CTTCTTCTCC	300
CGGGTGGTGC CCTCAGACAC ATACCAGGCC CAGGCCATGG TGGATATTGT CCGAGCCCTC	360
AAGTGGAACT ATGTGTCCAC ACTGGCCTCA GAGGGCAGCT ACGGTGAGAG TGGTGTGGAG	420
GCCTTTATCC AGAAGTCCCG AGAGAACGGA GGTGTGTGCA TTGCCCAGTC GGTGAAGATT	480
CCACGGGAAC CCAAGACGGG GGAGTTCGAC AAGATCATCA AACGCCTACT GGAAACATCC	540
AATGCCAGGG GTATCATCAT CTTTGCCAAC GAGGATGACA TCAGGAGGGT GTTGGAGGCA	600
GCTCGCAGGG CCAACCAGAC CGGCCACTTC TTTTGGATGG GTTCTGATAG CTGGGGCTCC	660
AAGAGTGCCC CTGTGCTGCG CCTTGAGGAG GTGGCCGAGG GCGCAGTCAC CATTCTCCCC	720
A A GA GGATGT CTGTTCGAGG GTTCGACCGA TACTTCTCCA GCCGCACGCT GGACAACAAC	780
AGGCGCAACA TCTGGTTTGC CGAGTTCTGG GAGGACAACT TCCATTGCAA GTTGAGCCGC	840
CACGCGCTCA AGAAGGGAAG CCACATCAAG AAGTGCACCA ACCGAGAGCG CATCGGGCAG	900
CACTCGCCT ATGAGCAGGA GGGGAAGGTG CAGTTCGTGA TTGACGCTGT GTACGCCATG	960
GGCCACGCGC TGCACGCCAT GCACCGTGAC CTGTGTCCCG GCCGCGTAGG ACTCTGCCCT	1020
CCCATGGACC CCGTGGATGG CACCCAGCTG CTTAAGTACA TCAGGAACGT CAACTTCTCA	1080
GCCATTGCGG GGAACCCTGT AACCTTCAAT GAGAACGGAG ACGCACCGGG GCGCTACGAC	1140
ATOTACCAGT ACCARCTGCG CAATGGCTCG GCCGAGTACA AGGTCATCGG CTCGTGGACA	1200
CACCACCTEC ACCTCAGAAT AGAGCGGATG CAGTGGCCAG GGAGTGGCCA GCAGCTGCCG	1260
COMPONENT GOAGTOTGCC CTGCCAGCCC GGGGAGCGAA AGAAGACTGT GAAGGGCATG	1320
CONTROTTECT GCCACTGCGA GCCCTGCACC GGGTACCAGT ACCAAGTGGA CCGCTACACC	. 1300
TGTAAGACCT GCCCCTACGA CATGCGGCCC ACAGAGAACC GCACGAGCTG CCAGCCCATC	1440
*Attnomicae ac-	

•				Programme and the contract of		A Company of the Comp	
	CCATCGTCA	AGTTGGAGTG	GGACTCGCCG	TGGGCCGTGC	TGCCCCTCTT	CCTGGCCGTG	1500
	TGGGCATCG	CCGCCACGCT	GTTCGTGGTG	GTCACGTTTG	TGCGCTACAA	CGATACCCCC	1560
,	TCGTCAAGG	CCTCGGGCCG	GGAGCTGAGC	TACGTGCTGC	TGGCGGGCAT	CTTTCTGTGC	1620
1	PACGCCACTA	CCTTCCTCAT	GATCGCAGAG	CCGGACCTGG	GGACCTGTTC	GCTCCGCCGC	1680
. 1	ATCTTCCTAG	GGCTCGGCAT	GAGCATCAGC	TACGCGGCCC	TGCTGACCAA	GACCAACCGC	1740
	ATTACCGCA	TCTTTGAGCA	GGGCAAACGG	TCGGTCAGTG	CCCCGCGTTT	CATCAGCCCG	1800
	CCTCCCAGC	TGGCCATCAC	CTTCATCCTC	ATCTCCCTGC	AGCTGCTCGG	CATCTGCGTG	1860
,		TGGACCCCTC	CCACTCGGTG	GTGGACTTCC	AGGACCAACG	GACACTTGAC	1920
	cccccanac idalicaida	CCAGGGGCGT	GCTCAAGTGC	: GACATCTCGG	ACCTGTCCCT	CATCTGCCTG	1980
	CCCCCT116	GCATGCTGCT	GATGGTCACG	TGTACTGTGT	ACGCCATCAA	GACCCGAGGC	2040
	cmccccch Gh	CCTTCAACGA	GGCCAAGCCC	ATCGGCTTCA	CCATGTACAC	CACCTGCATT	2100
	emerico CTICO	COTTOATCC	CATCTTTTT	r ggcacctcac	AGTCAGCCGA	CAAGCTGTAC	2160
	STCTGGCTGG	CCACACTGAG	GGTCTCCGT	AGTCTGAGC	CTTCAGTGTC	CCTGGGGATG	2220
	ATCCAGACAL	CONTRACTOR	CATCATCCT	C TTCCATATT	TTCCATTCT	CTCCTGGCCT	2280
	CICIACATGO		CCCTGCCC	C TCCTCCCTG!	A GCTGCCCCA	r cccccccatc	2340
	TCCCCTGCC		TCGATCTCA	T TTCCTACCA	CCTTCCCCC	r ACTTGGCTTC	2400
		•		- · .	•		2426
	CTCCACCAA	TCTTTCACC	n 64114			i.	

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Arg Asp Ser L u Ile Ser Ile Arg Asp Glu Lys Asp Gly Leu Asn Arg

Cys

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 16 amino acids (B) TYPE: amino acids (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22: Asp Arg Leu Leu Arg Lys Leu Arg Glu Arg Leu Pro Lys Ala Arg V. 2) INFORMATION FOR SEQ ID NO:23: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 16 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23: Glu Glu Val Trp Phe Asp Glu Lys Gly Asp Ala Pro Gly Arg Tyr J 1 5 10 15 (2) INFORMATION FOR SEQ ID NO:24: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 15 amino acids (B) TYPE: amino acids (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24: Glu Phe Val Tyr Glu Arg Glu Gly Asn Thr Glu Glu Asp Glu Leu 1 5 10 (2) INFORMATION FOR SEQ ID NO:25: (A) LENGTH: 17 amino acids (B) TYPE: amin acid (C) STRANDEDNESS: single							-							
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22: Asp Arg Leu Leu Arg Lys Leu Arg Glu Arg Leu Pro Lys Ala Arg V 1	1		(B) (C)	TYPE	: aml NDEDN	no ac: ESS: :	ia single	ds		X+			· ·	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22: Asp Arg Leu Leu Arg Lys Leu Arg Glu Arg Leu Pro Lys Ala Arg V 1		(ii) N	OLE	CULE	TYPE:	pept	ide			•	• .	• .	÷	
Asp Arg Leu Leu Arg Lys Leu Arg Glu Arg Leu Pro Lys Ala Arg V 1	-	•	•						e		. **		•	0
INFORMATION FOR SEQ ID NO:23: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 16 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23: Glu Glu Val Trp Phe Asp Glu Lys Gly Asp Ala Pro Gly Arg Tyr 1 (2) INFORMATION FOR SEQ ID NO:24: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 15 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24: Glu Phe Val Tyr Glu Arg Glu Gly Asn Thr Glu Glu Asp Glu Leu 1 (2) INFORMATION FOR SEQ ID NO:25: (A) LENGTH: 17 amino acids (B) TYPE: amin acid (C) STRANDEDNESS: single (A) LENGTH: 17 amino acids (B) TYPE: amin acid (C) STRANDEDNESS: single		(xi) \$	SEQU!	ENCE	DESCR	EIPTIO	N: SEQ	ID NO	:22:	•		. ·.	• - 1	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 16 amino acids (B) TYPE: amino acids (C) STRAMDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYFE: peptide (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23: Glu Glu Val Trp Phe Asp Glu Lys Gly Asp Ala Pro Gly Arg Tyr 1 (2) INFORMATION FOR SEQ ID NO:24: (1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 15 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24: Glu Phe Val Tyr Glu Arg Glu Gly Asn Thr Glu Glu Asp Glu Leu 1 5 10 (2) INFORMATION FOR SEQ ID NO:25: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 17 amino acids (B) TYPE: amino acids (C) STRANDENESS: single		Asp 1	Arg	Leu I	eu Ar 5	g Lys	Leu A	rg Glu	1 Arg	Leu l	Pro :	ra i	lla .	Arg v
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 16 amino acids (B) TYPE: amino acids (C) STRAMDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYFE: peptide (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23: Glu Glu Val Trp Phe Asp Glu Lys Gly Asp Ala Pro Gly Arg Tyr 1 (2) INFORMATION FOR SEQ ID NO:24: (1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 15 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24: Glu Phe Val Tyr Glu Arg Glu Gly Asn Thr Glu Glu Asp Glu Leu 1 5 10 (2) INFORMATION FOR SEQ ID NO:25: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 17 amino acids (B) TYPE: amino acids (C) STRANDENESS: single		2		,				m 8	*			, i	P-	
(A) LENGTH: 16 amino acid (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23: Glu Glu Val Trp Phe Asp Glu Lys Gly Asp Ala Pro Gly Arg Tyr 1 1 5 10 15 (2) INFORMATION FOR SEQ ID NO:24: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 15 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24: Glu Phe Val Tyr Glu Arg Glu Gly Asn Thr Glu Glu Asp Glu Leu 1 5 10 (2) INFORMATION FOR SEQ ID NO:25: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 17 amino acids (B) TYPE: amin acid (C) STRANDEDNESS: single	2)	INFOR	MATI	ON FO	RSE	Q ID N	0:23:		· ,					
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23: Glu Glu Val Trp Phe Asp Glu Lys Gly Asp Ala Pro Gly Arg Tyr 1 1 5 10 15 (2) INFORMATION FOR SEQ ID NO:24: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 15 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24: Glu Phe Val Tyr Glu Arg Glu Gly Asn Thr Glu Glu Asp Glu Leu 1 5 10 (2) INFORMATION FOR SEQ ID NO:25: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 17 amino acids (B) TYPE: amino acids (C) SEQUENCE CHARACTERISTICS: single		(i)	(A) (B)	LENG TYPI STRA	TH: E: am: NDED	ino ac NESS:	id single	Lus		***	\$ \$			
Glu Glu Val Trp Phe Asp Glu Lys Gly Asp Ala Pro Gly Arg Tyr 1 (2) INFORMATION FOR SEQ ID NO:24: (i) SEQUENCE CHARACTERISTICS:	• •	(ii)	MOLE	CULE	TYPE	: pept	ide		*				•	=
(2) INFORMATION FOR SEQ ID NO:24: (1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 15 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24: Glu Phe Val Tyr Glu Arg Glu Gly Asn Thr Glu Glu Asp Glu Leu 1 5 10 (2) INFORMATION FOR SEQ ID NO:25: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 17 amino acids (B) TYPE: amin acid (C) STRANDEDNESS: single		(xi)	SEQU	jence	DESC	RIPTIC	ON: SE	2 ID N	0:23:			 Gl.:	2	Tur 1
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 15 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24: Glu Phe Val Tyr Glu Arg Glu Gly Asn Thr Glu Glu Asp Glu Leu 1 5 10 15 (2) INFORMATION FOR SEQ ID NO:25: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 17 amino acids (B) TYPE: amin acid (C) STRANDEDNESS: single	 	Glu 1	Glu	Val '	Trp P 5	he As	p Glu	Lys Gl	y Asp 10	Ala	b.co	GTÅ	ALY	15
(A) LENGTH: 15 amino acid (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24: Glu Phe Val Tyr Glu Arg Glu Gly Asn Thr Glu Glu Asp Glu Leu 10 15 (2) INFORMATION FOR SEQ ID NO:25: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 17 amino acids (B) TYPE: amin acid (C) STPANDEDNESS: single	(2)	INFOR	TAMS	ON F	OR SE	QID	NO:24:		ø.	• •	1			*
(ii) MOLECULE TYPE: peptide (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24: Glu Phe Val Tyr Glu Arg Glu Gly Asn Thr Glu Glu Asp Glu Leu 10 15 (2) INFORMATION FOR SEQ ID NO:25: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 17 amino acids (B) TYPE: amin acid (C) STPANDEDNESS: single		(1)	(A (B) LEN) TYP) STR	GTH: E: al ANDEI	ino a NESS:	cid singl	140		*			:17 70	
Glu Phe Val Tyr Glu Arg Glu Gly Asn Thr Glu Glu Asp Glu Leu 1	•	(ii)								•			• • • •	
Glu Phe Val Tyr Glu Arg Glu Gly Asn Thr Glu Glu Asp Glu Leu 1		*	•	•					٠	٠.		0		
(2) INFORMATION FOR SEQ ID NO:25: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 17 amino acids (B) TYPE: amin acid (C) STRANDEDNESS: single	•	(xi)	SEQ	UENCE	DES	CRIPTI	ON: SI	Q ID	NO:24:			Agn	Glu	. Leu
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 17 amino acids (B) TYPE: amin acid (C) STRANDEDNESS: single	~		Phe	Val	Tyr	Glu Ar 5	g Glu	Gly A	sn Thi 10	. GIN	, GIA	vah		15
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 17 amino acids (B) TYPE: amin acid (C) STRANDEDNESS: single	(2)	INFO	RMAT	ION I	FOR S	EQ ID	NO: 25		* *	•	Φ			•
			SEC ()	UENCI) LEI) TYI	CHAI NGTH: PE: a	RACTER 17 ar min & DNESS:	RISTICATION ACIDA SING	s: cids		:				

- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Pro Glu Arg Lys Cys Cys Glu Ile Arg Glu Gln Tyr Gly Ile Gln Arg

Val

- (2) INFORMATION FOR SEQ ID NO:26:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Ile Gly Pro Gly Ser Ser Ser Val Ala Ile Gln Val Gln Asn Leu Leu
10 15

- (2) INFORMATION FOR SEQ ID NO:27:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 15 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Ile Ala Tyr Ser Ala Thr Ser Ile Asp Leu Ser Asp Lys Thr Leu
15

- (2) INFORMATION FOR SEQ ID NO:28:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 amino acids
 - (B) TYPE: amin acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Pro Glu Phe Ser Pro Ser Ser Gln Cys Pro Ser Ala His Ala Gln Leu 1 15

- (2) INFORMATION FOR SEQ ID NO:30:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 15 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Asp Lys Ile Ile Lys Arg Leu Leu Glu Thr Ser Asn Ala Arg Gly
1 5

- (2) INFORMATION FOR SEQ ID NO:31:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Val Asn Phe Ser Gly Ile Ala Gly Asn Pro Val Thr Phe Asn Glu Asn 10 15

- (2) INFORMATION FOR SEQ ID NO:32:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Gly Glu Ala Lys Ser Glu Leu Cys Glu Asn Leu Glu Thr Pro Ala Leu
10 15

- (2) INFORMATION FOR SEQ ID NO:33:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Pro Ala Arg Leu Ala Leu Pro Ala Asn Asp Thr Glu Phe Ser Ala Trp

Val

25

WHAT IS CLAIMED IS:

- 1. An isolated mammalian G protein-coupled glutamate receptor or a fragment thereof.
- 2. The G protein-coupled glutamate receptor of claim 1, which is substantially pure.
- 3. The G protein-coupled glutamate receptor of claim
 10 1, which is human or rodent.
 - 4. An antiserum obtained from an animal immunized with the G protein-coupled glutamate receptor of claim 1.
- 5. A monoclonal antibody which specifically binds to the G protein-coupled glutamate receptor of claim 1.
- 6. The G protein-coupled glutamate receptor of claim
 1, which binds glutamate or quisqualate and thereby activates
 phospholipase C or stimulates inositol phospholipid metabolism
 in a vertebrate cell.
 - 7. A recombinantly produced polypeptide having the activity of a mammalian G protein-coupled glutamate receptor.
 - 8. The polypeptide of claim 7, which has the activity of a human or rodent mammalian G protein-coupled glutamate receptor.
- 9. An isolated and purified polynucleotide molecule which codes for a mammalian G protein-coupled glutamate receptor or a fragment thereof.
- 10. The polynucleotide of claim 9, which is a genomic DNA sequence, a cDNA sequence, or an RNA antisense sequence.

- 11. The polynucleotide of claim 9, which codes for human or rodent G protein-c upled glutamate receptor.
- 12. The polynucleotide of claim 9, which encodes a polypeptide displaying mammalian G protein-coupled glutamate receptor activity.
- 13. The polynucleotide of claim 9, which is substantially the sequence of Fig. 5, Fig. 7, Fig. 8 or Fig. 9.

10

14. A probe which comprises an oligonucleotide capable of specifically hybridizing with a gene which encodes a mammalian G protein-coupled glutamate receptor or a fragment thereof.

15

- 15. The probe of claim 14, which comprises from about 40 to about 60 nucleotides in length.
- 16. The probe of claim 15, which is labeled to provide a detectable signal.
 - 17. A DNA construct comprising the following operably linked elements:
 - a transcriptional promoter;
- a DNA sequence encoding a mammalian G protein-coupled glutamate receptor or a fragment thereof; and
 - a transcriptional terminator.
 - 18. The DNA construct of claim 17, wherein the DNA sequence encodes a human or rodent G protein-coupled glutamate receptor polypeptide.
 - 19. The DNA construct of claim 17, wherein the DNA sequence encoding the mammalian G protein-coupled glutamate receptor is substantially the sequence of Fig. 5, Fig. 7, Fig. 8 or Fig. 9.

- 20. A cultured eukaryotic cell transformed or transf cted with a DNA construct which comprises th following operably linked elements:
 - a transcriptional promoter;
- a DNA sequence encoding a mammalian G protein-coupled glutamate receptor or a fragment thereof; and a transcriptional terminator.
- 21. The eukaryotic cell of claim 20, which is a mammalian cell.
 - 22. The eukaryotic cell of claim 20, which does not express endogenous G protein-coupled glutamate receptors.
- 23. The eukaryotic cell line of claim 20, wherein the DNA sequence encodes a human or rodent G protein-coupled glutamate receptor polypeptide.
- 24. The eukaryotic cell line of claim 21, wherein the G protein-coupled glutamate receptor polypeptide encoded by the DNA sequence is coupled to G protein in a mammalian cell.
- 25. The DNA construct of claim 20, wherein the DNA sequence encoding the mammalian G protein-coupled glutamate receptor is substantially the sequence of Fig. 5, Fig. 7, Fig. 8 or Fig. 9.
 - 26. A method for producing a mammalian G proteincoupled glutamate receptor, which comprises:
- growing eukaryotic cells transformed or transfected with a DNA construct which comprises a DNA sequence coding for the expression of the G protein-coupled glutamate receptor, and isolating the receptor from the cells.
- 27. The method of claim 26, wherein the c 11s are cultured mammalian cells.

WO 92/10583

- 28. The method of claim 26, wherein the glutamate receptor is human or rodent.
- 29. The method of claim 26, wherein the glutamate receptor is isolated by immunoaffinity purification.
- 30. The method of claim 26, wherein the G protein-coupled glutamate receptor is not coupled to protein G in the eukaryotic cells.

10

5

- 31. A method for determining the presence of a mammalian G protein-coupled glutamate receptor in a biological sample, which comprises incubating the sample with a monospecific antibody which specifically binds to the receptor under conditions sufficient for immune complex formation and determining therefrom the presence of the immune complexes.
- 32. The method of claim 31, wherein the monospecific antibody is a monoclonal antibody or a purified antiserum.

20

15 ,

- 33. The method of claim 32, wherein the monospecific antibody is labeled.
- 34. A method for identifying a compound which alters
 25 G protein-coupled glutamate receptor mediated-metabolism, which
 comprises incubating the compound with eukaryotic cells which
 express recombinant mammalian G protein-coupled glutamate
 receptor and determining therefrom the effect of said compound
 on receptor-mediated metabolism in the cells.

- 35. The method of claim 34, wherein the compound is incubated with the receptor and ligand.
- 36. The method of claim 35, wherein the ligand is glutamate or quisqualate.

PCT/US91/09422

- 37. The method of claim 34, wherein the eukaryotic c ll expresses a human or rodent G protein-coupled glutamate receptor.
- 38. The method of claim 37, wherein inositol phospholipid metabolism in the eukaryotic cell is monitored for alteration by the compound.

WO 92/10583 PCT/US91/09422

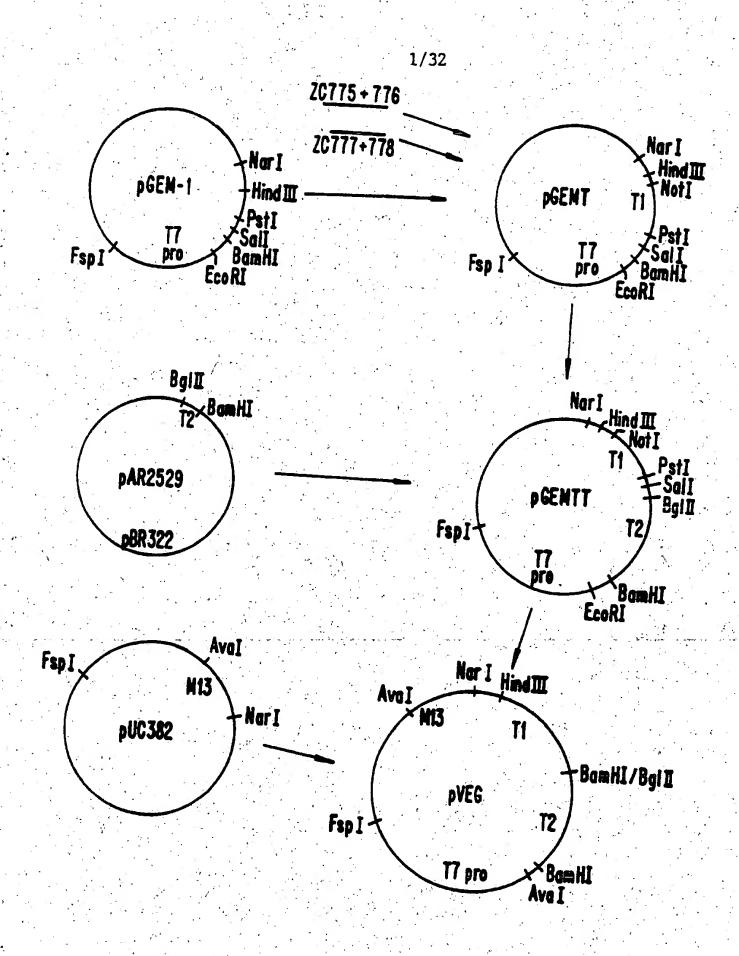


FIG. IA.

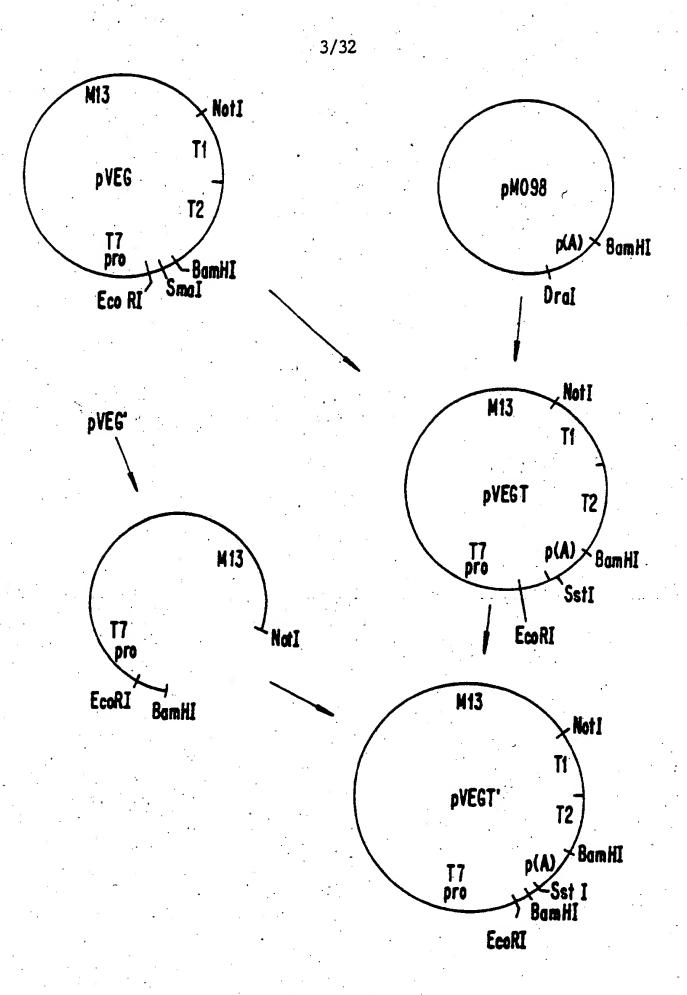
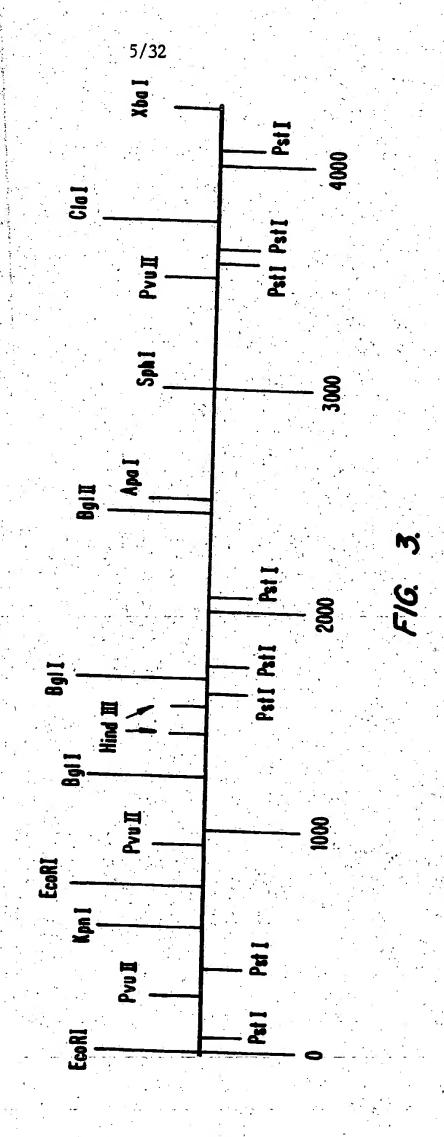
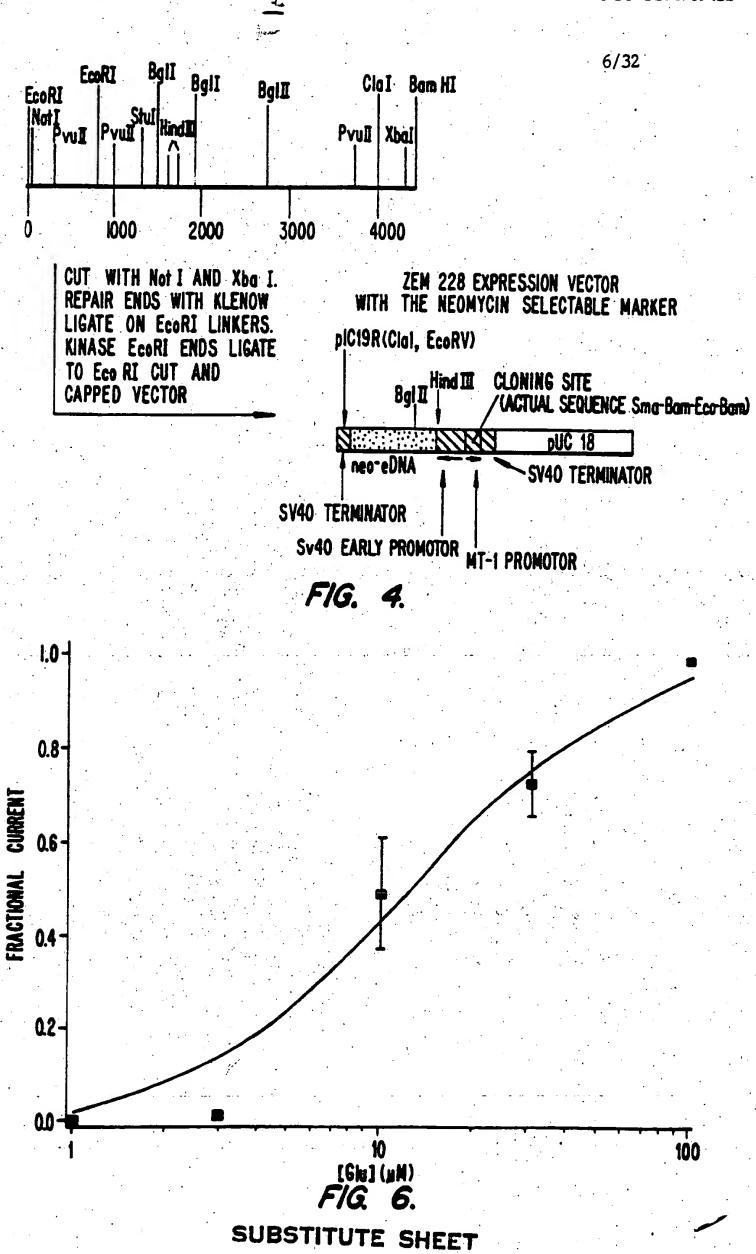


FIG. IC.

WO 92/10583 PCT/US91/09422



SUBSTITUTE SHEET



7/32 CCGAGAACGG CTGCAGTCCT CTGACCTGAG ACCAATAGCT GTGTCTACCC GGACTCAGCG TCCAGCTCAC CGCCACTAAC GCGCCGCGCA TTGGACACCT GATCCACACA CCTTCGGGCA 180 CCAGTGAAAA ACCGCGACTT GATTTTCTGG AAGAACGCCC CCAGGGTGTG GGAGCGGTCG TGGAGGACCA GCAGGAGGAA GCGGAGGGGA GAGGGCAGT AGTGGAGGCA GAGAAAGCGT TGAACCAGCT GTGTTGGCCG AAGGCACGAA ACGGCAAAAG GCAGCGGTGA GCATCTGTGT GGTTCCCGCT GGGAACCTGC AGGCAGGACC GGCGTGGGAA CGTGGCTGGC CCGCGGTGGA CCGCGTCTTC GCCACA ATG GTC CGG CTC CTC TTG ATT TTC TTC CCA ATG Met Val Arg Leu Leu Ile Phe Phe Pro Met ATC TIT TIG GAG ATG TCC ATT TIG CCC AGG ATG CCT GAC AGA AAA GTA Ile Phe Leu Glu Met Ser Ile Leu Pro Arg Met Pro Asp Arg Lys Val TTG CTG GCA GGT GCC TCG TCC CAG CGC TCC GTG GCG AGA ATG GAC GGA Leu Leu Ala Gly Ala Ser Ser Gln Arg Ser Val Ala Arg Met Asp Gly GAT GTC ATC ATC GGA GCC CTC TTC TCA GTC CAT CAC CAG CCT CCA GCC Asp Val Ile Ile Gly Ala Leu Phe Ser Val His His Gln Pro Pro Ala GAG AAG GTA CCC GAA AGG AAG TGT GGG GAG ATC AGG GAA CAG TAT GGT Glu Lys Val Pro Glu Arg Lys Cys Gly Glu Ile Arg Glu Gln Tyr Gly ATC CAG AGG GTG GAG GCC ATG TTC CAC ACG TTG GAT AAG ATT AAC GCG Ile Gln Arg Val Glu Ala Met Phe His Thr Leu Asp Lys Ile Asn Ala GAC CCG GTG CTC CTG CCC AAC ATC ACT CTG GGC AGT GAG ATC CGG GAC Asp Pro Val Leu Leu Pro Asn Ile Thr Leu Gly Ser Glu Ile Arg Asp 100 TCC TGC TGG CAC TCT TCA GTG GCT CTC GAA CAG AGC ATC GAA TTC ATC Ser Cys Trp His Ser Ser Val Ala Leu Glu Gln Ser Ile Glu Phe Ile

FIG. 5A.

				•					•		•				
AGA Arg	GAC Asp 125	TCC Ser	CTG Leu	ATT	TCC Ser	ATC	CGA	GAT Asp	GAG Glu	AAG Lys	GAT Asp 135	GGG Gly	CTG Leu	AAC Asn	793 CGA Arg
TGC Cys 140	CTG Leu	CCT Pro	GAT Asp	GGC G1y	CAG Gln 145	ACC Thr	CTG Leu	CCC	CCT Pro	GGC Gly 150	AGG Arg	ACT	AAG Lys	AAG Lys	841 CCT Pro 155
ATT	GCT Ala	GGA Gly	GTG Val	ATC Ile 160	GGC Gly	CCT Pro	GGC Gly	TCC Ser	AGC Ser 165	TCT Ser	GTG Val	GCC Ala	ATT	CAA Gln 170	889 GTC Val
CAG Gln	AAT Asn	CTT Leu	CTC Leu 175	CAG G1n	CTG Leu	TTC Phe	GAC Asp	ATC Ile 180	CCA	CAG G1n	ATC Ile	GCC Ala	TAT Tyr 185	TCT Ser	937 GCC Ala
ACA Thr	AGC Ser	ATA Ile 190	GAC Asp	CTG Leu	AGT Ser	GAC Asp	AAA Lys 195	ACT	TTG Leu	TAC Tyr	AAA Lys	TAC Tyr 200	TTC Phe	CTG Leu	985 AGG Arg
GTG Val	GTC Val 205	CCT Pro	TCT Ser	GAC Asp	ACT Thr	Leu	Gln	GCA Ala	AGG Arg	GCG Ala	ATG Met 215	CTC Leu	GAC Asp	ATA	O33 GTC Val
AAG Lys 220	CGT Arg	TAC Tyr	AAC Asn	TGG Trp	ACC Thr 225	TAT Tyr	GTC Val	TCA Ser	GCA Ala	GTC Val 230	CAC His	ACA Thr	GAA G1u	GGG	081 AAT Asn 235
TAC Tyr	GGC Gly	GAG G1u	AGT Ser	GGA Gly 240	ATG Met	GAT Asp	GCT Ala	TTC Phe	AAA Lys 245	GAA G1u	CTG Leu	GCT Ala	Ala	CAG Gln	129 GAA Glu
GGC Gly	CTC Leu	TGC Cys	ATC Ile 255	GCA Ala	CAC His	TCG Ser	GAC Asp	AAA Lys 260	ATC Ile	TAC Tyr	AGC Ser	AAT Asn	GCT Ala 265	GGC	177 GAG Glu
														CCC	
														GGC	
	TGC 140 ATT GTG AAG AAG AAG AAG AAG AAG AAG AAG	Arg Asp 125 TGC CTG Cys 140 ATT GCT Leu ATT Ala CAG AAT AGC Thr Ser GTG GTC Val 205 AAG CGT Val 205 AAG CGT CYal 205 AAG CGT Arg 220 TAC GGC CTC Gly GCC AGG AGG AGC AGG AGG AGG AGG AGG AGG	Arg Asp Ser 125 TGC CTG CCT Cys Leu Pro 140 ATT GCT GGA ATT CTT GIn Asn Leu ACA AGC ATA The 190 GTG GTC CCT Val Val Pro 205 AAG CGT TAC Tyr Gly Glu GGC CTC TGC GIy Glu GGC CTC TGC GIy Glu GCC AGG GTT ACT Cys Arg Val AAG AGC TTT ACT Cys Arg CTT ACT Cys Arg CTT Cys CTT Cys	TGC CTG CCT GAT Cys Leu Pro Asp 140 ATT GCT GGA GTG Ile Ala Gly Val CAG AAT CTT CTC Gln Asn Leu 175 ACA AGC ATA GAC Thr Ser 190 GTG GTC CCT TCT Val Val Pro Ser 205 AAG CGT TAC AAC Lys Arg Tyr Asn 220 TAC GGC GAG AGT Tyr Gly Glu Ser GGC CTC TGC ATC Gly Ceu Cys 255 AAG AGC TTT GAC Asp 270 GCC AGG GTT GTG ASp 270	Arg Asp Ser Leu Ile TGC CTG CCT GAT GGC Cys Leu Pro Asp Gly 140 ATT GCT GGA GTG ATC Ile Ala Gly Val Ile 160 CAG AAT CTT CTC CAG Gln Asn Leu Leu Gln 175 ACA AGC ATA GAC CTG Thr Ser Ile Asp Leu 190 GTG GTC CCT TCT GAC Val Val Pro Ser Asp 205 AAG CGT TAC AAC TGG Lys Arg Tyr Asn Trp 220 TAC GGC GAG AGT GGA Tyr Gly Glu Ser Gly 240 GGC CTC TGC ATC GCA Gly Leu Cys Ile Ala 255 AAG AGC TTT GAC CGG Lys Ser Phe Asp Arg 270 GCC AGG GTT GTG GTC Ala Arg Val Val	Arg Asp Ser Leu Ile Ser TGC CTG CCT GAT GGC CAG Cys Leu Pro Asp Gly Gln 140 ATT GCT GGA GTG ATC GGC Ile Ala Gly Val Ile Gly 160 CAG AAT CTT CTC CAG CTG Gln Asn Leu Leu Gln Leu 175 ACA AGC ATA GAC CTG AGT Thr Ser Ile Asp Leu Ser 190 GTG GTC CCT TCT GAC ACT Val Val Pro Ser Asp Thr 205 AAG CGT TAC AAC TGG ACC Lys Arg Tyr Asn Trp Thr 225 TAC GGC GAG AGT GGA ATG Tyr Gly Glu Ser Gly Met 240 GGC CTC TGC ATC GCA CAC Gly Leu Cys Ile Ala His 255 AAG AGC TTT GAC CGG CTC Lys Ser Phe Asp Arg Leu 270 GCC AGG GTT GTG GTC TGC Ala Arg Val Val Val Cys	AGA GAC TCC CTG ATT TCC ATC Arg Asp Ser Leu Ile Ser Ile 130 TGC CTG CCT GAT GGC CAG ACC Cys Leu Pro Asp Gly Gln Thr 140 ATT GCT GGA GTG ATC GGC CCT Ile Ala Gly Val Ile Gly Pro 160 CAG AAT CTT CTC CAG CTG TTC Gln Asn Leu Leu Gln Leu Phe 175 ACA AGC ATA GAC CTG AGT GAC Thr Ser Ile Asp Leu Ser Asp 190 GTG GTC CCT TCT GAC ACT TTG Val Val Pro Ser Asp Thr Leu 205 AAG CGT TAC AAC TGG ACC TAT Lys Arg Tyr Asn Trp Thr Tyr 220 TAC GGC GAG AGT GGA ATG GAT Tyr Gly Glu Ser Gly Met Asp 240 GGC CTC TGC ATC GCA CAC TCG GIy Leu Cys Ile Ala His Ser 255 AAG AGC TTT GAC CGG CTC CTG Lys Ser Phe Asp Arg Leu Leu Cys Phe GCC AGG GTT GTG GTC TGC TTC Ala Arg Val Val Val Cys Phe	TGC CTG CCT GAT GGC CAG ACC CTG Cys Leu Pro Asp Gly Gln Thr Leu 140 ATT GCT GGA GTG ATC GGC CCT GGC Ile Ala Gly Val Ile Gly Pro Gly 160 CAG AAT CTT CTC CAG CTG TTC GAC Gln Asn Leu Leu Gln Leu Phe Asp 175 ACA AGC ATA GAC CTG AGT GAC AAA Thr Ser Ile Asp Leu Ser Asp Lys 190 GTG GTC CCT TCT GAC ACT TTG CAG Val Val Pro Ser Asp Thr Leu Gln 210 AAG CGT TAC AAC TGG ACC TAT GTC Lys Arg Tyr Asn Trp Thr Tyr Val 225 TAC GGC GAG AGT GGA ATG GAT GCT Tyr Gly Glu Ser Gly Met Asp Ala 240 GGC CTC TGC ATC GCA CAC TCG GAC Gly Leu Cys Ile Ala His Ser Asp 275 AAG AGC TTT GAC CGG CTC CTG CGT Lys Ser Phe Asp Arg Leu Leu Arg 275 GCC AGG GTT GTG GTC TGC TTC TGC Ala Arg Val Val Val Cys Phe Cys	AGA GAC TCC CTG ATT TCC ATC CGA GAT Arg Asp Ser Leu Ile Ser Ile Arg Asp 125 TGC CTG CCT GAT GGC CAG ACC CTG CCC Cys Leu Pro Asp Gly Gin Thr Leu Pro 140 ATT GCT GGA GTG ATC GGC CCT GGC TCC Ile Ala Gly Val Ile Gly Pro Gly Ser 160 CAG AAT CTT CTC CAG CTG TTC GAC ATC Gln Asn Leu Leu Gln Leu Phe Asp Ile 175 ACA AGC ATA GAC CTG AGT GAC AAA ACT Thr Ser Ile Asp Leu Ser Asp Lys Thr 190 GTG GTC CCT TCT GAC ACT TTG CAG GCA CAT Val Val Pro Ser Asp Thr Leu Gln Ala 205 AAG CGT TAC AAC TGG ACC TAT GTC TCA Leu Gln Ala 220 TAC GGC GAG AGT GGA ATG GAT GCT TTC Tyr Gly Glu Ser Gly Met Asp Ala Phe 240 GGC CTC TGC ATC GCA CAC TCG GAC AAA CAC TYr Gly Glu Ser Gly Met Asp Ala Phe 240 AAG AGC TTT GAC CGG CTC CTG CGT AAA Lys Ser Phe Asp Arg Leu Leu Arg Lys 275 GCC AGG GTT GTG GTC TGC TTC TGC GAG AAA Arg Val Val Val Cys Phe Cys Glu	AGA GAC TCC CTG ATT TCC ATC CGA GAT GAG Arg Asp Ser Leu Ile Ser Ile Arg Asp Glu 125 TGC CTG CCT GAT GGC CAG ACC CTG CCC CCT Cys Leu Pro Asp Gly Gln Thr Leu Pro Pro 140 ATT GCT GGA GTG ATC GGC CCT GGC TCC AGC Ile Ala Gly Val Ile Gly Pro Gly Ser Ser 165 CAG AAT CTT CTC CAG CTG TTC GAC ATC CCA Gln Asn Leu Leu Gln Leu Phe Asp Ile Pro 175 ACA AGC ATA GAC CTG AGT GAC AAA ACT ITG Thr Ser Ile Asp Leu Ser Asp Lys Thr Leu 190 GTG GTC CCT TCT GAC ACT TTG CAG GCA AGG Val Pro Ser Asp Thr Leu Gln Ala Arg 205 AAG CGT TAC AAC TGG ACC TAT GTC TCA GCA Lys Arg Tyr Asn Trp Ihr Tyr Val Ser Ala 220 TAC GGC GAG AGT GGA ATG GAT GCT TTC AAA TTYr Gly Glu Ser Gly Met Asp Ala Phe Lys 240 GCC CTC TGC ATC GCA CAC TCG GAC AAA ATC CGI Leu Cys Ile Ala His Ser Asp Lys Ile 270 AAG AGC TTT GAC CGG CTC CTG CGT AAA CTC Lys Ser Phe Asp Arg Leu Leu Arg Lys Leu 270 GCC AGG GTT GTG GTC TGC TTC TGC GAG GGC Ala Arg Val Val Val Cys Phe Cys Glu Gly	AGA GAC TCC CTG ATT TCC ATC CGA GAT GAG AAG AAG ASP Ser Leu Ile Ser Ile Arg Asp Glu Lys 125 TGC CTG CCT GAT GGC CAG ACC CTG CCC CCT GGC Cys Leu Pro Asp Gly Gln Thr Leu Pro Pro Gly 140 ATT GCT GGA GTG ATC GGC CCT GGC TCC AGC TCT Ile Ala Gly Val Ile Gly Pro Gly Ser Ser Ser 165 CAG AAT CTT CTC CAG CTG TTC GAC ATC CCA CAG GIn Asn Leu Leu Gln Leu Phe Asp Ile Pro Gln 180 ACA AGC ATA GAC CTG AGT GAC AAA ACT TTG TAC Thr Ser Ile Asp Leu Ser Asp Lys Thr Leu Tyr 190 GTG GTC CCT TCT GAC ACT TTG CAG GCA AGG GCG Val Val Pro Ser Asp Thr Leu Gln Ala Arg Ala 210 AAG CGT TAC AAC TGG ACC TAT GTC TCA GCA GTC Val Val Pro Ser Asp Thr Tyr Val Ser Ala 230 TAC GGC GAG AGT GGA ATG GAT GAT GCT TTC AAA GAA TYR Gly Glu Ser Gly Met Asp Ala Phe Lys Glu 245 GGC CTC TGC ATC GCA CAC TCG GAC AAA ACT TAC GTC GIV Leu Cys Ile Ala His Ser Asp Lys Ile Tyr 260 GGC CTC TGC ATC GCA CAC TCG GAC AAA ATC TAC GGC GTC Lys Ser Phe Asp Arg Leu Leu Arg 275 GCC AGG GTT GTG GTC TGC TTC TGC GAG GGC ATG AAA ATC TAC GAG AAG ATG CTC CTG CTG CTG GAG GGC ATG AAA ATC TAC GAG AAG AGC TTT GAC CAC CTC CTG CTG CTG AAA CTC CGG Lys Ser Phe Asp Arg Leu Leu Arg 275 GCC AGG GTT GTG GTC TGC TTC TGC GAG GGC ATG AAA ATC Val Val Val Val Val Cys Phe Cys Glu Gly Met	AGA GAC TCC CTG ATT TCC ATC CGA GAT GAG AAG GAT ASP Ser Leu Ile Ser Ile Arg Asp Glu Lys Asp 125 TGC CTG CCT GAT GGC CAG ACC CTG CCC CCT GGC AGG Cys Leu Pro Asp Gly Gln Thr Leu Pro Pro Gly Arg 140 ATT GCT GGA GTG ATC GGC CCT GGC TCC AGC TCT GTG Ile Ala Gly Val Ile Gly Pro Gly Ser Ser Ser Val 160 CAG AAT CTT CTC CAG CTG TTC GAC ATC CCA CAG ATC GIn Asn Leu Leu Gln Leu Phe Asp 11e Pro Gln Ile 175 ACA AGC ATA GAC CTG AGT GAC AAA ACT TTG TAC AAA Thr Ser Ile Asp Leu Ser Asp Lys Thr Leu Tyr Lys 190 GTG GTC CCT TCT GAC ACT TTG CAG GCA AGG GCG ATG Wal Val Pro Ser Asp Thr Leu Gln Ala Arg Ala Met 205 AAG CGT TAC AAC TGG ACC TAT GTC TCA GCA GTC CAC Lys Arg Tyr Asn Trp Thr Tyr Val Ser Ala Wal His 220 TAC GGC GAG AGT GGA ATG GAT GCT TTC AAA GAA CTG Tyr Gly Glu Ser Gly Met Asp Ala Phe Lys Glu Leu 245 AAG AGC TT GC ATC GCA CAC TCG GAC AAA ATC TAC AGC GIy Leu Cys Ile Ala His Ser Asp Lys Ile Tyr Ser AAG AGG AGG GTG ATG ACA AGG AGG GTT TYR Ser Phe Asp Arg Leu Leu Arg Lys Leu Arg Glu GCC AGG GTT GTG GTG TTC TGC GAG GGC ATG ACA ATG ATG Wal Val Val Val Cys Phe Cys Glu Gly Met Thr	AGA GAC TCC CTG ATT TCC ATC CGA GAT GAG AAG GAT GGG Arg Asp Ser Leu IIe Ser IIe Arg Asp Glu Lys Asp Gly 135 TGC CTG CCT GAT GGC CAG ACC CTG CCC CCT GGC AGG ACT Cys Leu Pro Asp Gly Gln Thr Leu Pro Pro Gly Arg Thr 140 ATT GCT GGA GTG ATC GGC CCT GGC TCC AGC TCT GTG GCC TIE AT GTG GGC GTG ATC GGC GTG ATC GGC GTG ATC GGC GTG ATC GGC GTG ASP ASP GTG ATC GGC GTG ATC GGC GTG ASP ASP GTG ASP GTG ASP GTG ATC GGC GTG ATC GGC GTG ASP ASP GTG ASP GTG ATC GGC GTG ATC GGC GTG ASP GTG ASP GTG ASP GTG ASP GTG ASP GTG ATC GGC GTG ATC GGC GTG ASP GTG ASP GTG ASP GTG ASP GTG ASP GTG AGC ATG GTG GTG GTC CCT TCT GAC ACT TTG CAG GCA AGG GCG ATG CTC Val Val Pro Ser Asp Thr Leu Gln Ala Arg Ala Met Leu 205 AAG CGT TAC AAC TGG ACC TAT GTC TCA GCA GGC ATG CTC Val Val Pro Ser Asp Trp Try Val Ser Ala Val His Thr 220 TAC GGC GAG AGT GGA ATG GAT GCT TTC AAA GAA CTG GCT Tyr Gly Glu Ser Gly Met Asp Ala Phe Lys Glu Leu Ala 240 GGC CTC TGC ATC GCA CAC TCG GAC AAA ATC TAC AGC AAT GTG GTG TYR GLY Glu Ser GTY ASP GTY ASP GTY ASP GTY GTY Ser ASP CASP GTY	AGA GAC TCC CTG ATT TCC ATC CGA GAT GAG AAG GAT GGG CTG Arg Asp Ser Leu Tie Ser Tie Arg Asp Glu Lys Asp Gly Leu 125 TGC CTG CCT GAT GGC CAG ACC CTG CCC CCT GGC AGG ACT AAG Cys Leu Pro Asp Gly Gln Thr Leu Pro Pro Gly Arg Thr Lys 140 ATT GCT GGA GTG ATC GGC CCT GGC TCC AGC TCT GTG GCC ATT TIE Ala Gly Val Tie Gly Pro Gly Ser Ser Ser Val Ala Tie 160 CAG AAT CTT CTC CAG CTG TTC GAC ATC CCA CAG ATC GCC TAT Gln Asn Leu Leu Gln Leu Phe Asp Tie Pro Gln Tie Ala Tyr 185 ACA AGC ATA GAC CTG AGT GAC AAA ACT TTG TAC AAA TAC TTC Thr Ser Tie Asp Leu Ser Asp Lys Thr Leu Tyr Lys Tyr Phe 190 GTG GTC CCT TCT GAC ACT TTG CAG GCA AGG GCG ATG CTC GAC Val Val Pro Ser Asp Thr Leu Gln Ala Arg Ala Met Leu Asp 205 AAG CGT TAC AAC TGG ACC TAT GTC TCA GCA GTC CAC ACA GAA Lys Arg Tyr Asn Trp Thr Tyr Val Ser Ala Val His Thr Glu 220 TAC GGC GAG AGT GGA ATG GAT GAT GCT TTC AAA ATC TTC GTY GTY GTY GTY AS ATG TYR GTY AS ATG TYR TYR Val Ser Ala Val Leu Ala Ala CTY GTY GTY GTY GTY AS ATG TYR TYR AS ATG TYR TYR Val Ser Asp Lys Glu Leu Ala Ala CTY GTY GTY GTY GTY AS ATG TYR GTY AS ATG TYR GTY AS ATG TYR GTY GTY AS ATG TYR GTY AS ATG TYR GTY GTY AS ATG TYR GTY GTY GTY GTY AS ATG TYR GTY AS ATG TYR GTY AS ATG TYR GTY AS ATG TYR GTY GTY AS ATG TYR GTY GTY AS ATG TYR GTY AS ATG TYR GTY AS ATG TYR GTY GTY AS ATG TYR GTY AS ATG	AGA GAC TCC CTG ATT TCC ATC CGA GAT GAG AAG GAT GGG CTG AAC ATG ASP Ser Leu Ile Ser Ile Arg Asp Glu Lys Asp Gly Leu Asn 135 TGC CTG CCT GAT GGC CAG ACC CTG CCC CCT GGC AGG ACT AAG AAG CYS Leu Pro Asp Gly Gln Thr Leu Pro Pro Gly Arg Thr Lys Lys 140 ATT GCT GGA GTG ATC GGC CCT GGC TCC AGC TCT GTG GCC ATT CAA Ile Ala Gly Val Ile Gly Pro Gly Ser Ser Ser Val Ala Ile Gln 160 CAG AAT CTT CTC CAG CTG TTC GAC ATC CCA CAG ATC GCC TAT TCT GIN Asn Leu Leu Gln Leu Phe Asp 180 ACA AGC ATA GAC CTG AGT GAC AAA ACT TTG TAC AAA TAC TTC CTG Thr Ser Ile Asp Leu Ser Asp Lys Thr Leu Tyr Lys Tyr Phe Leu 190 GTG GTC CCT TCT GAC ACT TTG CAG GCA AGG GCG ATG CTC GAC ATA Val Val Pro Ser Asp Thr Leu Gln Ala Arg Ala Met Leu Asp Ile 210 AAG CGT TAC AAC TGG ACC TAT GTC TCA GCA GTC CAC ACA GAG GGG Lys Arg Tyr Asn Trp Thr Tyr Val Ser Ala Val His Thr Glu Gly 220 TAC GGC GAG AGT GGA ATG GAT GCT TTC AAA ATC TTC AAA GAA CTG GCT GCC CAG Tyr Gly Glu Ser Gly Met Asp Ala Phe Lys Glu Leu Ala Ala Gly 245 GGC CTC TGC ATC GCA CAC TCG GAC AAA ATC TC CGG GAG CGG CTT CCC GGC Gly Leu Cys Ile Ala His Ser Asp Lys Ile Tyr Ser Asn Ala Gly 250 GGC CTC TGC ATC GCA CAC TCG GAC AAA ACT TC CGG GAG CGG CTT CCC CAG GTY GTY GIV Glu Ser Asp Arg Leu Leu Arg Lys Ile Tyr Ser Asn Ala Gly 250 GCC AGG GTT GGG GTC TGC TCC CTG CGT AAA CTC CGG GAG CGG CTT CCC CAG ATA Arg Val Val Val Val Cys Phe Cys Glu Gly Met Thr Val Arg Gly Arg Leu Pro 270

FIG. 5B.

1321 CTG AGT GCC ATG CGC CGC CTG GGC GTG GGC GAG TTC TCA CTC ATT Leu Ser Ala Met Arg Arg Leu Gly Val Val Gly Glu Phe Ser Leu Ile 1369 GGA AGT GAT GGA TGG GCA GAC AGA GAT GAA GTC ATC GAA GGC TAT GAG Gly Ser Asp Gly Trp Ala Asp Arg Asp Glu Val Ile Glu Gly Tyr Glu 325 GTG GAA GCC AAC GGA GGG ATC ACA ATA AAG CTT CAG TCT CCA GAG GTC Val Glu Ala Asn Gly Gly Ile Thr Ile Lys Leu Gln Ser Pro Glu Val 340 AGG TCA TTT GAT GAC TAC TTC CTG AAG CTG AGG CTG GAC ACC AAC ACA Arg Ser Phe Asp Asp Tyr Phe Leu Lys Leu Arg Leu Asp Thr Asn Thr 350 AGG AAT CCT TGG TTC CCT GAG TTC TGG CAA CAT CGC TTC CAG TGT CGC Arg Asn Pro Trp Phe Pro Glu Phe Trp Gln His Arg Phe Gln Cys Arg 370 **365** CTA CCT GGA CAC CTC TTG GAA AAC CCC AAC TTT AAG AAA GTG TGC ACA Leu Pro Gly His Leu Leu Glu Asn Pro Asn Phe Lys Lys Val Cys Thr 390 395 GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AAA ATG GGA Gly Asn Glu Ser Leu Glu Glu Asn Tyr Val Gln Asp Ser Lys Met Gly TTT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CAG AAC ATG Phe Val Ile Asn Ala Ile Tyr Ala Met Ala His Gly Leu Gln Asn Met 415 420 · CAC CAT GCT CTG TGT CCC GGC CAT GTG GGC CTG TGT GAT GCT ATG AAA His His Ala Leu Cys Pro Gly His Val Gly Leu Cys Asp Ala Met Lys 430 CCC ATT GAT GGC AGG AAG CTC CTG GAT TTC CTC ATC AAA TCC TCT TTT Pro Ile Asp Gly Arg Lys Leu Leu Asp Phe Leu Ile Lys Ser Ser Phe 445 450 GTC GGA GTG TCT GGA GAG GAG GTG TGG TTC GAT GAG AAG GGG GAT GCT Val Gly Val Ser Gly Glu Glu Val Trp Phe Asp Glu Lys Gly Asp Ala 465 460 475

FIG. 5C.

	•						. 15	•	*.							-
	CCC	GGA Gly	AGG Arg	TAT Tyr	GAC Asp 480	ATT Ile	ATG Met	AAT Asn	CTG	CAG Gln 485	TAC Tyr	ACA Thr	GAA Glu	GCT Ala	AAT	1849 CGC Arg
	TAT Tyr	GAC Asp	TAT Tyr	GTC Val 495	CAC His	GTG Val	GGG Gly	ACC Thr	TGG Trp 500	CAT His	GAA Glu	GGA Gly	GTG Val	CTG Leu 505	AAT	1897 ATT Ile
	GAT Asp	GAT Asp	TAC Tyr 510	AAA Lys	ATC Ile	CAG G1n	ATG Met	AAC Asn 515	AAA Lys	AGC Ser	GGA Gly	ATG Met	GTA Val 520	CGA Arg	TCT	1945 GTG Val
		AGT Ser 525			TGC Cys	TTA Leu	AAG Lys 530	GGT Gly	CAG G1n	ATT Ile	AAG Lys	GTC Val 535	ATA Ile	CGG Arg	AAA	1993 GGA Gly
	GAA G1u 540	GTG Val	AGC Ser	TGC Cys	TGC Cys	TGG Trp 545	ÅTC Ile	TGC Cys	ACG Thr	GCC Ala	TGC Cys 550	AAA Lys	GAG G1u	AAT Asn	GAG	2041 TTT Phe 555
4	GTG Val	CAG G1n	GAC Asp	GAG G1u	TTC Phe 560	ACC Thr	TGC Cys	AGA Arg	GCC Ala	TGT Cys 565	GAC Asp	CTG Leu	GGG Gly	TGG Trp	TGG	2089 CCC Pro
	AAC Asn	GCA Ala	GAG G1u	CTC Leu 575	ACA Thr	GGC Gly	TGT Cys	GAG Glu	CCC Pro 580	ATT Ile	CCT Pro	GTC Val		TAT Tyr 585	CTT	
	TGG Trp	AGT Ser	GAC Asp 590	ATA Ile	GAA G1u	TCT Ser	ATC Ile	ATA Ile 595	Ala	ATC Ile	GCC Ala	TTT Phe		TGC Cys	CTG	
	ATC Ile			ACG Thr	CTG Leu	TTT Phe	GTC Val 610	ACC Thr	CTC Leu	ATC Ile	TTC Phe	GTT Val 615		TAC Tyr	CGG	
	ACA Thr 620	CCC Pro	GTG Val	GTC Val	AAA Lys	TCC Ser 625	TCC Ser	AGT Ser	AGG Arg	GAG Glu	CTC Leu 630	TGC Cys	TAT Tyr		ATT	281 CTG Leu 635
	GCT Ala	GGT Gly	ATT	TTC Phe	CTC Leu 640	GGC Gly	TAT Tyr	GTG Val	TGC Cys	CCT Pro 645	TTC Phe	ACC Thr	CTC Leu	ATC Ile	GCC	AAA Lys
		•								\						

FIG. 5D.

		•			•					1			•			
	CCT Pro	ACT Thr	ACC Thr	ACA Thr 655	TCC Ser	TGC Cys	TAC Tyr	CTC Leu	CAG Gln 660	CGC Arg	CTC Leu	CTA Leu	GTT Val	GGC Gly 665	CTC Leu	2377 TCT Ser
r															ATT Ile	
	CGC Arg	ATC Ile 685	CTG Leu	GCT Ala	GGC Gly	AGC Ser	AAG Lys 690	AAG Lys	AAG Lys	ATC Ile	TGC Cys	ACC Thr 695	CGG Arg	AAG Lys	CCC Pro	2473 AGA Arg
	TTC Phe 700	ATG Met	AGC Ser	GCT Ala	TGG Trp	GCC Ala 705	CAA G1n	GTG Val	ATC Ile	ATA Ile	GCC Ala 710	TCC Ser	ATT	CTG Leu	ATT	2521 AGT Ser 715
												ATG Met				
	CCC Pro	ATT	TTG Leu	TCC Ser 735	TAC Tyr	CCG Pro	AGT Ser	ATC Ile	AAG Lys 740	GAA Glu	GTC Val	TAC Tyr	CTT Leu	ATC Ile 745	TGC ² Cys	AAT ASN
	ACC Thr	AGC Ser	AAC Asn 750	CTG Leu	GGT Gly	GTA Val	GTG Val	GCC Ala 755	CCT	GTG Val	GGT Gly	TAC Tyr	AAT Asn 760	GGA Gly	CTC Leu	2665 CTC Leu
	ATC Ile	ATG Met 765	AGC Ser	TGT Cys	ACC Thr	TAC Tyr	TAT Tyr 770	GCC Ala	TTC Phe	AAG Lys	ACC Thr	CGC Arg 775	AAC Asn	GTG Val	CCG Pro	GCC Ala
	AAC Asn 780	TTC Phe	AAT Asn	GAG Glu	GCT Ala	AAA Lys 785	TAC Tyr	ATC Ile	GCC Ala	TTC Phe	ACC Thr 790	ATG Met	TAC Tyr	ACT Thr	ACC	761 TGC Cys 795
	ATC Ile	ATC Ile	TGG Trp	CTG Leu	GCT Ala 800	TTC Phe	GTT Val	CCC Pro	ATT	TAC Tyr 805	TTT Phe	GGG Gly	AGC Ser	AAC Asn		
	ATC Ile	ATC Ile	ACT Thr	ACC Thr 815	TGC Cys	TTC Phe	GCG Ala	GTG Val	AGC Ser 820	CTC Leu	AGT Ser	GTG Val	ACG Thr	GTG Val 825	GCC Ala	2857 CTG Leu

FIG. 5E.

			-					12/3	-				÷			
	GGG Gly	TGC Cys	ATG Met 830	TTT Phe	ACT Thr	CCG Pro	AAG Lys	ATG Met 835	TAC Tyr	ATC Ile	ATC Ile	ATT Ile	GCC Ala 840	AAA Lys	CCT	2905 GAG Glu
	Arg									TCT Ser				CGC Arg	ATG	
	GTC Val 860													CTC Leu	AAC	3001 ATT Ile 875
											GCC Ala			AAC Asn		
	TCT Ser													AAG Lys 905		
	CAC His	GTG Val									AAG Lys			GAG G1u		
														TAC Tyr	CAA	
														ACC Thr	CTT	
- +	AAT Asn	GTG Val	GAA G1 u	GAA G1u	GAG G1u 960	GAC Asp	AAT Asn	ACC Thr	CCT	TCT Ser 965	GCT Ala	CAC His	TTC Phe	AGC Ser		
	AGC Ser	AGC Ser	CCT Pro	TCT Ser 975	ATG Met	GTG Val	GTG Val	CAC His	CGA Arg 980	CGC Arg	GGG Gly	CCA Pro	CCC Pro	GTG Val 985		
	ACA Thr	CCA Pro	CCT Pro 990	CTG Leu	CCA Pro	CCC Pro	CAT	CTG Leu 995	ACC Thr	GCA Ala	GAA G1u	GAG G1u	ACC Thr 1000	CCC Pro	CTG	385 TTC Phe

FIG. 5F.

WO 92/10583 PCT/US91/09422 13/32 CTG GCT GAT ICC GTC ATC CCC AAG GGC TTG CCT CCT CCC CCG CAG Leu Ala Asp Ser Val Ile Pro Lys Gly Leu Pro Pro Pro Leu Pro Gln 1005 CAG CAG CCA CAG CCG CCC CCT CAG CAG CCC CCG CAG CAG CCC AAG Gin Gin Pro Gin Gin Pro Pro Pro Gin Gin Pro Pro Gin Gin Pro Lys 1025 1035 TCC CTG ATG GAC CAG CTG CAA GGC GTA GTC ACC AAC TTC GGT TCG GGG Ser Leu Met Asp Gln Leu Gln Gly Val Val Thr Asn Phe Gly Ser Gly 1040 1045 ATT CCA GAT TTC CAT GCG GTG CTG GCA GGC CCG GGG ACA CCA GGA AAC Ile Pro Asp Phe His Ala Val Leu Ala Gly Pro Gly Thr Pro Gly Asn 1055 1060 AGC CTG CGC TCT CTG TAC CCG CCC CCG CCT CCG CCG CAA CAC CTG CAG Ser Leu Arg Ser Leu Tyr Pro Pro Pro Pro Pro Gln His Leu Gln 1070 1075 ATG CTG CCC CTG CAC CTG AGC ACC TTC CAG GAG GAG TCC ATC TCC CCT Met Leu Pro Leu His Leu Ser Thr Phe Gln Glu Glu Ser Ile Ser Pro 1085 CCT GGG GAG GAC ATC GAT GAC AGT GAG AGA TTC AAG CTC CTG CAG Pro Gly Glu Asp Ile Asp Asp Ser Glu Arg Phe Lys Leu Leu Gln 1110 1115 GAG TTC GTG TAC GAG CGC GAA GGG AAC ACC GAA GAA GAT GAA TTG GAA Glu Phe Val Tyr Glu Arg Glu Gly Asn Thr Glu Glu Asp Glu Leu Glu 1120 1125 GAG GAG GAG CTG CCC ACA GCC AGC AAG CTG ACC CCT GAG GAT TCT Glu Glu Glu Asp Leu Pro Thr Ala Ser Lys Leu Thr Pro Glu Asp Ser 1140 CCT GCC CTG ACG CCT CCT TCT CCT TTC CGA GAT TCC GTG GCC TCT GGC Pro Ala Leu Thr Pro Pro Ser Pro Phe Arg Asp Ser Val Ala Ser Gly

AGC TCA GTG CCC AGT TCC CCC GTA TCT GAG TCG GTC CTC TGC ACC CCT Ser Ser Val Pro Ser Ser Pro Val Ser Glu Ser Val Leu Cys Thr Pro 1165

1155

FIG. 5G.

CCA AAT GTA ACC TAC GCC TCT GTC ATT CTG AGG GAC TAC AAG CAA AGC Pro Asn Val Thr Tyr Ala Ser Val Ile Leu Arg Asp Tyr Lys Gln Ser 1180

TCT TCC ACC CTG TAGTGTGTGT GTGTGTGTGG GGGCGGGGG AGTGCGCATG Ser Ser Thr Leu

FIG. 5H.

TAAGAATTTT A	ATAAATACTC	TGGGAATTTT	ATTGGTGATG	CCTTTGTGTC	TACAGGGCA
ACGTTCCAGA	GAGCTCTGGT	GTGAAGTGAT	GGGGGACTTG	TGGCTAGAGA	AGCTTTTCA/
TGGCCTTAAA (CTCTGGGTCC	TGCTTGAGAG	AGGTCTGAGG	TTCTCAACAT	180 CAGAGCAGA
CTTCCACCAA G	GCTTTCAGAA	TGCTAAGCCC	CCACTTCTCA	ACACTTAGTG	240 CTCTGATCG
TGCCTGCGAA C	CCGAGAACGG	CTGCAGTCCT	CTGACCTGAG	ACCAATAGCT	300 GTGTCTACCO
GGACTCAGCG T	CCAGCTCAC	CGCCACTAAC	GCGCCGCGCA	TTGGACACCT	360 GATCCACACA
CCTTCGGGCA C	CAGTGAAAA	ACCGCGACTT	GATTTTCTGG	AAGAACGCCC	420 CCAGGGTGT
GGAGCGGTCG T	GGAGGACCA	GCAGGAGGAA	GCGGAGGGA	GAGGGGCAGT	480 AGTGGAGGCA
GAGAAAGCGT T	GAACCAGCT	GTGTTGGCCG	AAGGCACGAA	ACGGCAAAAG	540 GCAGCGGTGA
GCATCTGTGT G	GTTCCCGCT	GGGAACCTGC	AGGCAGGACC	GGCGTGGGAA	600 CGTGGCTGGC
CCGCGGTGGA C	CGCGTCTTC		GTC CGG CTC Val Arg Leu		
CCA ATG ATC Pro Met Ile 10	TTT TTG GA Phe Leu Gl 1	G ATG TCC u Met Ser	ATT TTG CCC Ile Leu Pro 20	AGG ATG CCT Arg Met Pro	GAC AGA Asp Arg 25
AAA GTA TTG Lys Val Leu	CTG GCA GG Leu Ala Gl 30	T GCC TCG y Ala Ser	TCC CAG CGC Ser Gln Arg 35	TCC GTG GCG Ser Val Ala	749 AGA ATG Arg Met 40
GAC GGA GAT Asp Gly Asp	GTC ATC ATC Val Ile Ile 45	C GGA GCC e Gly Ala	CTC TTC TCA Leu Phe Ser 50	GTC CAT CAC Val His His 55	797 CAG CCT Gln Pro
CCA GCC GAG Pro Ala Glu 60					
TAT GGT ATC	CAG AGG GT	G GAG GCC		ACG TTG GAT	AAG ATT

FIG. 7B.

17/32

1469 GGC GAG AAG AGC TTT GAC CGG CTC CTG CGT AAA CTC CGG GAG CGG CTT Gly Glu Lys Ser Phe Asp Arg Leu Leu Arg Lys Leu Arg Glu Arg Leu 270 280 CCC AAG GCC AGG GTT GTG GTC TGC TTC TGC GAG GGC ATG ACA GTG CGG Pro Lys Ala Arg Val Val Cys Phe Cys Glu Gly Met Thr Val Arg 295 1565 GGC TTA CTG AGT GCC ATG CGC CGC CTG GGC GTC GTG GGC GAG TTC TCA Gly Leu Leu Ser Ala Met Arg Arg Leu Gly Val Val Gly Glu Phe Ser 1613 CTC ATT GGA AGT GAT GGA TGG GCA GAC AGA GAT GAA GTC ATC GAA GGC Leu Ile Gly Ser Asp Gly Trp Ala Asp Arg Asp Glu Val Ile Glu Gly **320**. 315 1661 TAT GAG GTG GAA GCC AAC GGA GGG ATC ACA ATA AAG CTT CAG TCT CCA Tyr Glu Val Glu Ala Asn Gly Gly Ile Thr Ile Lys Leu Gln Ser Pro 330 335 340 345 1709 GAG GTC AGG TCA TTT GAT GAC TAC TTC CTG AAG CTG AGG CTG GAC ACC Glu Val Arg Ser Phe Asp Asp Tyr Phe Leu Lys Leu Arg Leu Asp Thr 350 360 1757 AAC ACA AGG AAT CCT TGG TTC CCT GAG TTC TGG CAA CAT CGC TTC CAG Asn Thr Arg Asn Pro Trp Phe Pro Glu Phe Trp Gln His Arg Phe Gln 365 1805 TGT CGC CTA CCT GGA CAC CTC TTG GAA AAC CCC AAC TTT AAG AAA GTG Cys Arg Leu Pro Gly His Leu Leu Glu Asn Pro Asn Phe Lys Lys Val 380 390 TGC ACA GGA AAT GAA AGC TTG GAA GAA AAC TAT GTC CAG GAC AGC AAA Cys Thr Gly Asn Glu Ser Leu Glu Glu Asn Tyr Val Gln Asp Ser Lys 395 400 405 1901 ATG GGA TIT GTC ATC AAT GCC ATC TAT GCC ATG GCA CAT GGG CTG CAG Met Gly Phe Val Ile Asn Ala Ile Tyr Ala Met Ala His Gly Leu Gln 425 420 415 410 1949 AAC ATG CAC CAT GCT CTG TGT CCC GGC CAT GTG GGC CTG TGT GAT GCT Asn Met His His Ala Leu Cys Pro Gly His Val Gly Leu Cys Asp Ala 440 435 430

FIG. 7C.

ATG AAA CCC ATT GAT GGC AGG AAG CTC CTG GAT TTC CTC ATC AAA TCC Met Lys Pro Ile Asp Gly Arg Lys Leu Leu Asp Phe Leu Ile Lys Ser 2045 TCT TTT GTC GGA GTG TCT GGA GAG GAG GTG TGG TTC GAT GAG AAG GGG Ser Phe Val Gly Val Ser Gly Glu Glu Val Trp Phe Asp Glu Lys Gly 465 2093 GAT GCT CCC GGA AGG TAT GAC ATT ATG AAT CTG CAG TAC ACA GAA GCT Asp Ala Pro Gly Arg Tyr Asp Ile Met Asn Leu Gln Tyr Thr Glu Ala AAT CGC TAT GAC TAT GTC CAC GTG GGG ACC TGG CAT GAA GGA GTG CTG Asn Arg Tyr Asp Tyr Val His Val Gly Thr Trp His Glu Gly Val Leu 490 495 505 2189 AAT ATT GAT GAT TAC AAA ATC CAG ATG AAC AAA AGC GGA ATG GTA CGA Asn Ile Asp Asp Tyr Lys Ile Gln Met Asn Lys Ser Gly Met Val Arg TCT GTG TGC AGT GAG CCT TGC TTA AAG GGT CAG ATT AAG GTC ATA CGG Ser Val Cys Ser Glu Pro Cys Leu Lys Gly Gln Ile Lys Val Ile Arg 525 2285 AAA GGA GAA GTG AGC TGC TGC TGC ATC TGC ACG GCC TGC AAA GAG AAT Lys Gly Glu Val Ser Cys Cys Trp Ile Cys Thr Ala Cys Lys Glu Asn 540 2333 GAG TTT GTG CAG GAC GAG TTC ACC TGC AGA GCC TGT GAC CTG GGG TGG Glu Phe Val Gln Asp Glu Phe Thr Cys Arg Ala Cys Asp Leu Gly Trp 555 560 2381 TGG CCC AAC GCA GAG CTC ACA GGC TGT GAG CCC ATT CCT GTC CGT TAT Trp Pro Asn Ala Glu Leu Thr Gly Cys Glu Pro Ile Pro Val Arg Tyr 570 585 580 CTT GAG TGG AGT GAC ATA GAA TCT ATC ATA GCC ATC GCC TTT TCT TGC Leu Glu Trp Ser Asp Ile Glu Ser Ile Ile Ala Ile Ala Phe Ser Cys **590** 595 600 2477 CTG GGC ATC CTC GTG ACG CTG TTT GTC ACC CTC ATC TTC GTT CTG TAC Leu Gly Ile Leu Val Thr Leu Phe Val Thr Leu Ile Phe Val Leu Tyr 605 615 610 2525 CGG GAC ACA CCC GTG GTC AAA TCC TCC AGT AGG GAG CTC TGC TAT ATC Arg Asp Thr Pro Val Val Lys Ser Ser Ser Arg Glu Leu Cys Tyr Ile

FIG. 7D.

620

625

630

19/32

ATT CTG GCT GGT ATT ITC CTC GGC IAT GTG TGC CCT TTC ACC CTC ATC Ile Leu Ala Gly Ile Phe Leu Gly Tyr Val Cys Pro Phe Thr Leu Ile GCC AAA CCT ACT ACC ACA TCC TGC TAC CTC CAG CGC CTC CTA GTT GGC Ala Lys Pro Thr Thr Ser Cys Tyr Leu Gln Arg Leu Leu Val Gly 655 2669 CTC TCT TCT GCC ATG TGC TAC TCT GCT TTA GTG ACC AAA ACC AAT CGT Leu Ser Ser Ala Met Cys Tyr Ser Ala Leu Val Thr Lys Thr Asn Arg 2717 ATT GCA CGC ATC CTG GCT GGC AGC AAG AAG ATC TGC ACC CGG AAG Ile Ala Arg Ile Leu Ala Gly Ser Lys Lys Lys Ile Cys Thr Arg Lys CCC AGA TTC ATG AGC GCT TGG GCC CAA GTG ATC ATA GCC TCC ATT CTG Pro Arg Phe Met Ser Ala Trp Ala Gln Val Ile Ile Ala Ser Ile Leu ATT AGT GTA CAG CTA ACA CTA GTG GTG ACC TTG ATC ATC ATG GAG CCT Ile Ser Val Gln Leu Thr Leu Val Val Thr Leu Ile Ile Met Glu Pro CCC ATG CCC ATT TTG TCC TAC CCG AGT ATC AAG GAA GTC TAC CTT ATC Pro Met Pro Ile Leu Ser Tyr Pro Ser Ile Lys Glu Val Tyr Leu Ile 730 TGC AAT ACC AGC AAC CTG GGT GTA GTG GCC CCT GTG GGT TAC AAT GGA Cys Asn Thr Ser Asn Leu Gly Val Val Ala Pro Val Gly Tyr Asn Gly 750 CTC CTC ATC ATG AGC TGT ACC TAC TAT GCC TTC AAG ACC CGC AAC GTG Leu Leu Ile Met Ser Cys Thr Tyr Tyr Ala Phe Lys Thr Arg Asn Val 765 CCG GCC AAC TTC AAT GAG GCT AAA TAC ATC GCC TTC ACC ATG TAC ACT Pro Ala Asn Phe Asn Glu Ala Lys Tyr Ile Ala Phe Thr Met Tyr Thr 780 785 790 ACC TGC ATC ATC TGG CTG GCT TTC GTT CCC ATT TAC TTT GGG AGC AAC Thr Cys Ile Ile Trp Leu Ala Phe Val Pro Ile Tyr Phe Gly Ser Asn 800 FIG. 7E.

TAC AAG ATC ATC ACT ACC TGC TTC GCG GTG AGC CTC AGT GTG ACG GTG Tyr Lys Ile Ile Thr Thr Cys Phe Ala Val Ser Leu Ser Val Thr Val 825 810 3149 GCC CTG GGG TGC ATG TTT ACT CCG AAG ATG TAC ATC ATC ATT GCC AAA Ala Leu Gly Cys Met Phe Thr Pro Lys Met Tyr Ile Ile Ala Lys CCT GAG AGG AAC GTC CGC AGT GCC TTC ACG ACC TCT GAT GTC CGC Pro Glu Arg Asn Val Arg Ser Ala Phe Thr Thr Ser Asp Val Val Arg 845 ATG CAC GTC GGT GAT GGC AAA CTG CCG TGC CGC TCC AAC ACC TTC CTC Met His Val Gly Asp Gly Lys Leu Pro Cys Arg Ser Asn Thr Phe Leu 860 AAC ATT TTC CGG AGA AAG AAG CCC GGG GCA GGG AAT GCC AAG AAG AGG Asn Ile Phe Arg Arg Lys Lys Pro Gly Ala Gly Asn Ala Lys Lys Arg CAG CCA GAA TTC TCG CCC AGC AGC CAG TGT CCG TCG GCA CAT GCG CAG Gln Pro Glu Phe Ser Pro Ser Ser Gln Cys Pro Ser Ala His Ala Gln 900 CTT TGAAAACCCC CACACTGCAG TGAATGTTTC TAACGGCAAG TCTGTGTCAT Leu GGTCTGAACC AGGTGGAAGA CAGGCGCCCA AGGGACAGCA CGTGTGGCAG CGCCTCTCTG TGCACGTGAA GACCAACGAG ACGGCCTGTA ACCAAACAGC CGTAATCAAA CCCCTCACTA AAAGTTACCA AGGCTCTGGC AAGAGCCTGA CCTTTTCAGA TGCCAGCACC AAGACCCTTT ACAATGTGGA AGAAGAGGAC AATACCCCTT CTGCTCACTT CAGCCCTCCC AGCAGCCCTT 3694 CTATGGTGGT GCACCGACGC GGGCCACCCG TGGCCACCAC ACCACCTCTG CCACCCCATC TGACCGCAGA AGAGACCCCC CTGTTCCTGG CTGATTCCGT CATCCCCAAG GGCTTGCCTC CTCCTCTCCC GCAGCAGCAG CCACAGCAGC CGCCCCCTCA GCAGCCCCCG CAGCAGCCCA AGTCCCTGAT GGACCAGCTG CAAGGCGTAG TCACCAACTT CGGTTCGGGG ATTCCAGATT FIG. 7F.

TCCATGCGGT GCTGGCAGGC CCGGGGACAC CAGGAAACAG CCTGCGCTCT CTGTACCCGC CCCCGCCTCC GCCGCAACAC CTGCAGATGC TGCCCCTGCA CCTGAGCACC TTCCAGGAGG AGTCCATCTC CCCTCCTGGG GAGGACATCG ATGATGACAG TGAGAGATTC AAGCTCCTGC AGGAGTTCGT GTACGAGCGC GAAGGGAACA CCGAAGAAGA TGAATTGGAA GAGGAGGĀĞG ACCTGCCCAC AGCCAGCAAG CTGACCCCTG AGGATTCTCC TGCCCTGACG CCTCCTTCTC CTTTCCGAGA TTCCGTGGCC TCTGGCAGCT CAGTGCCCAG TTCCCCCGTA TCTGAGTCGG TCCTCTGCAC CCCTCCAAAT GTAACCTACG CCTCTGTCAT TCTGAGGGAC TACAAGCAAA CCAGAGATGC CAAGGAGTGT CAACCCTTCC AGAAATGTGT AGAAAGCAGG GTGAGGGATG AGGAGAGGGA CGATGCCAAC TGAACAGTGG TCCTGGCCAG GATTGTGACT CTTGAATTAT TCAAAAACCT TCTCTAGAAA GAAAGGGAAT TATGACAAAG CACAATTCCA TATGGTATGT TTTGCTCAAT CGTGCATACA TATATCTGCC CACACTCCCG TGGTAAAACT AGAAGCGAĀG CAGGCCCTGC GATGGTGCCA ACTGAATCCT AAGTTCATCA TCCTAGTGAG CAGATGGAGA 4834 GAGGGCAGGA GGCGAGAGGG CAGGAGGCGG GGGTAGGTTC GGACAACAGC TCCCATCTCA 4894 GACCTTGACT GTGCTGAGTC TTCAGACTCC TGGACTAAGG AAGACCCGGG GACTGACCTT ATGAGGGTCC CTTTCCACTG CTGTGATCCA TTGCCAGCCT GTAGTCACCC GGGATAAAGG CACAGTAACC TTTTGCATTC CTGTGATTCC CTGTGTTTAA GGAAAAGGAA AGTATGĂĞCA FIG. 7G

WO 92/10583 PCT/US91/09422

22/32

AAGCTATCAC CAAAAAGAGC GCCATTAGAA GTTACGGGGG AGAAAAAAAG AGAAGCAAGA
TGATATATAA GCACAGGGCC TTGAACAAGG TGAGCGTGCT TCACAGATTC CGTATTAATG
TACAGATACT TTTGGAGAGG AGAAAGATAA CAAGGAGTGT CAGGCCGTTT GTGAACTCAC
TTGCACTGTG CCAACCAGGT TCTCCGCTGC CCTTCAGCAA AA

FIG. 7H.

CCCGGGCTCC CGGCAGTGCG AGCAGCTAAG GGCTGGCCGC CGCCTCCCTG AGCTCCCCG GMGCAGCCGA CCCCTGGTCG CGGCGTTCAC CTCGCCGATG CGCGGTTGGT AGGAGTGACC GGAGCCATTC TCTCCTCGTT GATAAGATTC CCTACCAGGA TAGGAGCCTA TCTCCCTTTY CACAGCAGGA CACAGAAATC TGGCCTTCAG TACTTTGGGA AAAGGATCTG AGACCTCCTG GAGCTCTGAC CACTGGCTGT CATCTGTGGC TCTGGCCTGT GTGGGCCACT GAGCTCTĂČŤ CAAACATTAA AGAGGAGGAG GGGAGATCTG TGGAATGGGC CACCCCGTTG GCCTGCTGCA TTACTGAACC TGCGCTGTCC ACACGTGCCC AGATCATGGG ACCCAGGGCC TGCTAGGGCT AGGAGCGGGG CCCAGTATTC ATGGGTCTCT AGGCCTTTCC GAA ATG TCC GGG AAG Met Ser Gly Lys GGA GGC TGG GCC TGG TGG GCC CGG CTG CCC CTC TGC CTA CTC CTC Gly Gly Trp Ala Trp Trp Ala Arg Leu Pro Leu Cys Leu Leu Leu AGC CTT TAT GCC CCC TGG GTG CCT TCA TCC TTG GGA AAG CCC AAG GGT Ser Leu Tyr Ala Pro Trp Val Pro Ser Ser Leu Gly Lys Pro Lys Gly CAC CCC CAC ATG AAC TCT ATC CGA ATT GAC GGG GAC ATC ACA CTG GGA His Pro His Met Asn Ser Ile Arg Ile Asp Gly Asp Ile Thr Leu Gly 40 GGC CTG TTT CCC GTC CAC GGC CGT GGC TCT GAG GGT AAG GCC TGC GGG Gly Leu Phe Pro Val His Gly Arg Gly Ser Glu Gly Lys Ala Cys Gly GAG CTG AAG AAG GAG AAA GGC ATC CAC CGC CTG GAG GCC ATG CTG TTT Glu Leu Lys Lys Glu Lys Gly Ile His Arg Leu Glu Ala Met Leu Phe GCC CTG GAC CGC ATC AAC AAT GAC CCG GAC CTA CTG CCC AAC ATC ACG Ala Leu Asp Arg Ile Asn Asn Asp Pro Asp Leu Leu Pro Asn Ile Thr 85 100

FIG. 8A.

TTG GGC GCC CGC ATT CTG GAC ACC TGC TCG AGG GAC ACC CAC GCC CTG Leu Gly Ala Arg Ile Leu Asp Thr Cys Ser Arg Asp Thr His Ala Leu GAG CAG ICA CTG ACC ITT GTG CGG GCG CTC ATC GAG AAG GAC GGC ACG Glu Gln Ser Leu Thr Phe Val Arg Ala Leu Ile Glu Lys Asp Gly Thr GAG GTC CGC TGC GGC AGG CGG GGC CCG CCC ATC ATC ACC AAG CCC GAA Glu Val Arg Cys Gly Arg Arg Gly Pro Pro Ile Ile Thr Lys Pro Glu CGA GTG GTG GGT GTC ATT GGA GCT TCG GGG AGC TCC GTC TCG ATC ATG Arg Val Val Gly Val Ile Gly Ala Ser Gly Ser Ser Val Ser Ile Met 150 1003 GTG GCC AAC ATC CTC CGC CTC TTC AAG ATC CCT CAG ATC AGC TAT GCC Val Ala Asn Ile Leu Arg Leu Phe Lys Ile Pro Gln Ile Ser Tyr Ala TCC ACG GCC CCT GAC TTG AGT GAC AAC AGC CGC TAT GAC TTC TTC TCC Ser Thr Ala Pro Asp Leu Ser Asp Asn Ser Arg Tyr Asp Phe Phe Ser CGG GTG GTG CCC TCA GAC ACA TAC CAG GCC CAG GCC ATG GTG GAT ATT Arg Val Val Pro Ser Asp Thr Tyr Gln Ala Gln Ala Met Val Asp Ile 200 205 GTC CGA GCC CTC AAG TGG AAC TAT GTG TCC ACA CTG GCC TCA GAG GGC Val Arg Ala Leu Lys Trp Asn Tyr Val Ser Thr Leu Ala Ser Glu Gly 215 AGC TAC GGT GAG AGT GGT GTG GAG GCC TTT ATC CAG AAG TCC CGA GAG Ser Tyr Gly Glu Ser Gly Val Glu Ala Phe Ile Gln Lys Ser Arg Glu 230 235 AAC GGA GGT GTG TGC ATT GCC CAG TCG GTG AAG ATT CCA CGG GAA CCC Asn Gly Gly Val Cys Ile Ala Gln Ser Val Lys Ile Pro Arg Glu Pro 245 250 260 AAG ACG GGG GAG TTC GAC AAG ATC AAA CGC CTA CTG GAA ACA TCC Lys Thr Gly Glu Phe Asp Lys Ile Ile Lys Arg Leu Leu Glu Thr Ser 265 270 275

FIG. 8B.

25/32 AAT GCC AGG GGT ATC ATC ATC TTT GCC AAC GAG GAT GAC ATC AGG AGG Asn Ala Arg Gly Ile Ile Ile Phe Ala Asn Glu Asp Asp Ile Arg Arg GTG TTG GAG GCA GCT CGC AGG GCC AAC CAG ACC GGC CAC TTC TTT TGG Val Leu Glu Ala Ala Arg Arg Ala Asn Gln Thr Gly His Phe Phe Trp ATG GGT TCT GAT AGC TGG GGC TCC AAG AGT GCC CCT GTG CTG CGC CTT Met Gly Ser Asp Ser Trp Gly Ser Lys Ser Ala Pro Val Leu Arg Leu 310 GAG GAG GTG GCC GAG GGC GCA GTC ACC ATT CTC CCC AAG AGG ATG TCT Glu Glu Val Ala Glu Gly Ala Val Thr Ile Leu Pro Lys Arg Met Ser GTT CGA GGG TTC GAC CGA TAC TTC TCC AGC CGC ACG CTG GAC AAC AAC Val Arg Gly Phe Asp Arg Tyr Phe Ser Ser Arg Thr Leu Asp Asn Asn AGG CGC AAC ATC IGG ITT GCC GAG ITC IGG GAG GAC AAC TTC CAT TGC Arg Arg Asn Ile Trp Phe Ala Glu Phe Trp Glu Asp Asn Phe His Cys 360 365 AAG TTG AGC CGC CAC GCG CTC AAG AAG GGA AGC CAC ATC AAG AAG TGC Lys Leu Ser Arg His Ala Leu Lys Lys Gly Ser His Ile Lys Lys Cys 375 380 ACC AAC CGA GAG CGC ATC GGG CAG GAC TCG GCC TAT GAG CAG GAG GGG Thr Asn Arg Glu Arg Ile Gly Gln Asp Ser Ala Tyr Glu Gln Glu Gly 390 400 390 AAG GTG CAG TTC GTG ATT GAC GCT GTG TAC GCC ATG GGC CAC GCG CTG Lys Val Gln Phe Val Ile Asp Ala Val Tyr Ala Met Gly His Ala Leu 405 CAC GCC ATG CAC CGT GAC CTG TGT CCC GGC CGC GTA GGA CTC TGC CCT His Ala Met His Arg Asp Leu Cys Pro Gly Arg Val Gly Leu Cys Pro CGC ATG GAC CCC GTG GAT GGC ACC CAG CTG CTT AAG TAC ATC AGG AAC Arg Met Asp Pro Val Asp Gly Thr Gln Leu Leu Lys Tyr Ile Arg Asn 445 450

FIG. 8C.

26/32 GTC AAC TTC TCA GGC ATT GCG GGG AAC CCT GTA ACC TTC AAT GAG AAC Val Asn Phe Ser Gly Ile Ala Gly Asn Pro Val Thr Phe Asn Glu Asn 460 GGA GAC GCA CCG GGG CGC TAC GAC ATC TAC CAG TAC CAA CTG CGC AAT Gly Asp Ala Pro Gly Arg Tyr Asp Ile Tyr Gln Tyr Gln Leu Arg Asn GGC TCG GCC GAG TAC AAG GTC ATC GGC TCG TGG ACA GAC CAC CTG CAC Gly Ser Ala Glu Tyr Lys Val Ile Gly Ser Trp Thr Asp His Leu His 2011 CTC AGA ATA GAG CGG ATG CAG TGG CCA GGG AGT GGC CAG CAG CTG CCG Leu Arg Ile Glu Arg Met Gln Trp Pro Gly Ser Gly Gln Gln Leu Pro CGC TCC ATC TGC AGT CTG CCC TGC CAG CCC GGG GAG CGA AAG AAG ACT Arg Ser Ile Cys Ser Leu Pro Cys Gln Pro Gly Glu Arg Lys Lys Thr GTG AAG GGC ATG GCT TGC TGC TGC CAC TGC GAG CCC TGC ACC GGG TAC Val Lys Gly Met Ala Cys Cys Irp His Cys Glu Pro Cys Thr Gly Tyr CAG TAC CAA GTG GAC CGC TAC ACC TGT AAG ACC TGC CCC TAC GAC ATG Gln Tyr Gln Val Asp Arg Tyr Thr Cys Lys Thr Cys Pro Tyr Asp Met 550 CGG CCC ACA GAG AAC CGC ACG AGC TGC CAG CCC ATC CCC ATC GTC AAG Arg Pro Thr Glu Asn Arg Thr Ser Cys Gln Pro Ile Pro Ile Val Lys .570 565 TTG GAG TGG GAC TCG CCG TGG GCC GTG CTG CCC CTC TTC CTG GCC GTG Leu Glu Trp Asp Ser Pro Trp Ala Val Leu Pro Leu Phe Leu Ala Val 590° 585 GTG GGC ATC GCC GCC ACG CTG TTC GTG GTG GTC ACG TTT GTG CGC TAC Val Gly Ile Ala Ala Thr Leu Phe Val Val Val Thr Phe Val Arg Tyr 605 600 AAC GAT ACC CCC ATC GTC AAG GCC TCG GGC CGG GAG CTG AGC TAC GTG Asn Asp Thr Pro Ile Val Lys Ala Ser Gly Arg Glu Leu Ser Tyr Val 620 615

FIG. 8D.

27/32 CTG CTG GCG GGC ATC ITT CTG TGC TAC GCC ACT ACC TTC CTC ATG ATC Leu Leu Ala Gly Ile Phe Leu Cys Tyr Ala Thr Thr Phe Leu Met Ile 635 640 GCA GAG CCG GAC CTG GGG ACC TGT TCG CTC CGC CGC ATC TTC CTA GGG Ala Glu Pro Asp Leu Gly Thr Cys Ser Leu Arg Arg Ile Phe Leu Gly 650. CTC GGC ATG AGC ATC AGC TAC GCG GCC CTG CTG ACC AAG ACC AAC CGC Leu Gly Met Ser Ile Ser Tyr Ala Ala Leu Leu Thr Lys Thr Asn Arg ATT TAC CGC ATC TTT GAG CAG GGC AAA CGG TCG GTC AGT GCC CCG CGT Ile Tyr Arg Ile Phe Glu Gln Gly Lys Arg Ser Val Ser Ala Pro Arg 680 685 TTC ATC AGC CCG GCC TCG CAG CTG GCC ATC ACC TTC ATC CTC ATC TCC Phe Ile Ser Pro Ala Ser Gln Leu Ala Ile Thr Phe Ile Leu Ile Ser 700 CTG CAG CTG CTC GGC ATC TGC GTG TGG TTC GTG GTG GAC CCC TCC CAC Leu Gln Leu Leu Gly Ile Cys Val Trp Phe Val Val Asp Pro Ser His 710 720 TCG GTG GTG GAC TTC CAG GAC CAA CGG ACA CTT GAC CCC CGC TTT GCC Ser Val Val Asp Phe Gln Asp Gln Arg Thr Leu Asp Pro Arg Phe Ala 730 AGG GGC GTG CTC AAG TGC GAC ATC TCG GAC CTG TCC CTC ATC TGC CTG Arg Gly Val Leu Lys Cys Asp Ile Ser Asp Leu Ser Leu Ile Cys Leu 745 CTG GGC TAC AGC ATG CTG CTG ATG GTC ACG TGT ACT GTG TAC GCC ATC Leu Gly Tyr Ser Met Leu Leu Met Val Thr Cys Thr Val Tyr Ala Ile 760 765 AAG ACC CGA GGC GTG CCC GAG ACC TTC AAC GAG GCC AAG CCC ATC GGC Lys Thr Arg Gly Val Pro Glu Thr Phe Asn Glu Ala Lys Pro Ile Gly 780 TTC ACC ATG TAC ACC ACC TGC ATT GTC TGG CTG GCC TTC ATC CCC ATC Phe Thr Met Tyr Thr Thr Cys Ile Val Trp Leu Ala Phe Ile Pro Ile 795 800 790

FIG. 8E.

2923 TTT TTT GGC ACC TCA CAG TCA GCC GAC AAG CTG TAC ATC CAG ACA ACC Phe Phe Gly Thr Ser Gln Ser Ala Asp Lys Leu Tyr Ile Gln Thr Thr **810**. 2971 ACA CTG ACG GTC TCC GTG AGT CTG AGC GCT TCA GTG TCC CTG GGG ATG Thr Leu Thr Val Ser Val Ser Leu Ser Ala Ser Val Ser Leu Gly Met 825 3019 CTC TAC ATG CCC AAA GTC TAC ATC ATC CTC TTC CAC CCG GAG CAG AAC Leu Tyr Met Pro Lys Val Tyr Ile Ile Leu Phe His Pro Glu Gln Asn 3067 GTG CCC AAG CGC AAG CGC AGT CTC AAA GCC GTG GTC ACC GCC GCC ACC Val Pro Lys Arg Lys Arg Ser Leu Lys Ala Val Val Thr Ala Ala Thr 860 ATG TCC AAC AAG TTC ACA CAG AAG GGC AAC TTC AGG CCC AAT GGG GAA Met Ser Asn Lys Phe Thr Gln Lys Gly Asn Phe Arg Pro Asn Gly Glu 875 GCC AAA TCA GAG CTG TGT GAG AAC CTG GAG ACC CCA GCG CTG GCT ACC Ala Lys Ser Glu Leu Cys Glu Asn Leu Glu Thr Pro Ala Leu Ala Thr 890 **895**. .

AAA CAG ACC TAC GTC ACC TAC ACC AAC CAT GCC ATC TAGCCGGGCC Lys Gln Thr Tyr Val Thr Tyr Thr Asn His Ala Ile
905
910

GCGGAGCCAA GCAGGCTAAG GAGCCACAAC CTCTGAGGAT GGCACATTGG GCCAGGGCCG
TTCCCGAGGG CCCTGCCGAT GTCTGCCCGC CTCCCGGGCA TCCACGAATG TGGCTTGGTG
CTGAGGACAG TAGAGACCCC GGCCATCACT GCTGGGCAAG CCGTGGTGGG CAACCAGAGG
AGGCCGAGTG GCTGGGGCAG TTCCAGGTTA TGCCACACAC AGGTCTTCCT TCTGGACCAC
TGTTGGCCCA GCCCCAAAGC ACAGGGGCTC GGTCTCCAGA GCCCAGCCCT GGCTTCCTCT
CCTTCCTCCT GCCTCCGTCT GTCCTGTGGG TGACCCCGGT TGGTCCCTGC CCCGTCTTTA
CGTTTCTCTT CCGTCTTTGC TCTGCATGTG TTGTCTGTTT GGGCCCTCTG CTTCCATATT

FIG. 8F.

TITICATTCT GCTCCTGGCC TICCCCTGCC ATCTGCCCTG CCCCCTGCCC CTCCTCCCTG

AGCTGCCCCA TCCCCGCCAT CATTITCTCT TCTGTTCCCC CTCGATCTCA TITCCTACCA

GCCTTCCCCC TACTTGGCTT CATCCACCAA CTCTTTCACC ACGTTGCAAA AGAGAAAAAA

AAAGGGGGGG GGGAATCACC CCCTACAAAA AAGCCCAAAC AAAAACTAAT CTTGAGTGTG

TTTCGAAGTG CTGCGTCCTC CTGGTGGCCT GTGTGTCCCT GTGCCTGCAG CCTGTCTGCC

CGCCCTACCC GTCTGCCGTG TGTCCTGCCC CCCCCGCCTG CCCGCCTTGC CCTTCCTGCT

AACGACACGG AGTTCAGTGC CTGGGTGTTT GGTGATGGTC TCTGATGTGT AGCATGTCTG

TTTTTATACC GAGAACATTT CTAATAAAGA TAAACACATG GTTTTGC

FIG. 8G.

CCCAACATCA CGTTGGGCGC CCGCATTCTG GACACCTGCT CGAGGGAEAC CCACGCCCTG GAGCAGTCAC TGACCTTTGT GCGGGCGCTC ATCGAGAAGG ACGGCACGGA GGTCCGCTGC 180 GGCAGGCGGG GCCCGCCCAT CATCACCAAG CCCGAACGAG TGGTGGGTGT CATTGGAGCT TCGGGGAGCT CCGTCTCGAT CATGGTGGCC AACATCCTCC GCCTCTTCAA GATCCCTCAG 300 ATCAGCTATG CCTCCACGGC CCCTGACTTG AGTGACAACA GCCGCTATGA CTTCTTCTCC CGGGTGGTGC CCTCAGACAC ATACCAGGCC CAGGCCATGG TGGATATTGT CCGAGCCCTC AAGTGGAACT ATGTGTCCAC ACTGGCCTCA GAGGGCAGCT ACGGTGAGAG TGGTGTGGÄG 480 GCCTTTATCC AGAAGTCCCG AGAGAACGGA GGTGTGTGCA TTGCCCAGTC GGTGAAGATT CCACGGGAAC CCAAGACGGG GGAGTTCGAC AAGATCATCA AACGCCTACT GGAAACATCC AATGCCAGGG GTATCATCAT CTTTGCCAAC GAGGATGACA TCAGGAGGGT GTTGGAGĞČĀ GCTCGCAGGG CCAACCAGAC CGGCCACTTC TTTTGGATGG GTTCTGATAG CTGGGGCTCC AAGAGTGCCC CTGTGCTGCG CCTTGAGGAG GTGGCCGAGG GCGCAGTCAC CATTCTCCCC AAGAGGATGT CTGTTCGAGG GTTCGACCGA TACTTCTCCA GCCGCACGCT GGACAACAAC AGGCGCAACA TCTGGTTTGC CGAGTTCTGG GAGGACAACT TCCATTGCAA GTTGAGCČGČ CACGCGCTCA AGAAGGGAAG CCACATCAAG AAGTGCACCA ACCGAGAGCG CATCGGGCAG GACTCGGCCT ATGAGCAGGA GGGGAAGGTG CAGTTCGTGA TTGACGCTGT GTACGCCĂTĞ GGCCACGCGC TGCACGCCAT GCACCGTGAC CTGTGTCCCG GCCGCGTAGG ACTCTGCCCT CGCATGGACC CCGTGGATGG CACCCAGCTG CTTAAGTACA TCAGGAACGT CAACTTCTCA 1140 GGCATTGCGG GGAACCCTGT AACCTTCAAT GAGAACGGAG ACGCACCGGG GCGCTACGAC

FIG. 9A.

SUBSTITUTE SHEET

31/32 1200 ATCTACCAGT ACCAACTGCG CAATGGCTCG GCCGAGTACA AGGTCATCGG CTCGTGGACA 1260 GACCACCTGC ACCTCAGAAT AGAGCGGATG CAGTGGCCAG GGAGTGGCCA GCAGCTGCCG 1320 CGCTCCATCT GCAGTCTGCC CTGCCAGCCC GGGGAGCGAA AGAAGACTGT GAAGGGCATG 1380 GCTTGCTGCT GGCACTGCGA GCCCTGCACC GGGTACCAGT ACCAAGTGGA CCGCTACACC TGTAAGACCT GCCCCTACGA CATGCGGCCC ACAGAGAACC GCACGAGCTG CCAGCCCATC CCCATCGTCA AGTTGGAGTG GGACTCGCCG TGGGCCGTGC TGCCCCTCTT CCTGGCCGTG GTGGCCATCG CCGCCACGCT GTTCGTGGTG GTCACGTTTG TGCGCTACAA CGATACCCCC 1620 ATCGTCAAGG CCTCGGGCCG GGAGCTGAGC TACGTGCTGC TGGCGGGCAT CTTTCTGTGC 1680 TACGCCACTA CCTTCCTCAT GATCGCAGAG CCGGACCTGG GGACCTGTTC GCTCCGCCGC 1740 ATCTTCCTAG GGCTCGGCAT GAGCATCAGC TACGCGGCCC TGCTGACCAA GACCAACCGC ATTTACCGCA TCTTTGAGCA GGGCAAACGG TCGGTCAGTG CCCCGCGTTT CATCAGCCCG GCCTCGCAGC TGGCCATCAC CTTCATCCTC ATCTCCCTGC AGCTGCTCGG CATCTGCGTG 1920 TGGTTCGTGG TGGACCCCTC CCACTCGGTG GTGGACTTCC AGGACCAACG GACACTTGAC 1980 CCCCGCTTTG CCAGGGGCGT GCTCAAGTGC GACATCTCGG ACCTGTCCCT CATCTGCCTG 2040 CTGGGCTACA GCATGCTGCT GATGGTCACG TGTACTGTGT ACGCCATCAA GACCCGAGGC 2100 GTGCCCGAGA CCTTCAACGA GGCCAAGCCC ATCGGCTTCA CCATGTACAC CACCTGCATT 2160 GTCTGGCTGG CCTTCATCCC CATCTTTTTT GGCACCTCAC AGTCAGCCGA CAAGCTGTĂC 2220 ATCCAGACAA CCACACTGAC GGTCTCCGTG AGTCTGAGCG CTTCAGTGTC CCTGGGGATG 2280 CTCTACATGC CCAAAGTCTA CATCATCCTC TTCCATATTT TTCCATTCTG CTCCTGGCCT

FIG. 9B.

WO 92/10583 PCT/US91/09422

32/32

TCCCCTGCCA TCTGCCCTGC CCCCTGCCCC TCCTCCCTGA GCTGCCCCAT CCCCGCCATC ATTTTCTCTT CTGTTCCCCC TCGATCTCAT TTCCTACCAG CCTTCCCCCT ACTTGGCTTC CTCCACCAAC TCTTTCACCA CGTTGC

FIG 9C

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US91/09422

						icate all) ³
1		ational Patent Classi se See Attached		oth National Classifica	ition and IPC	
US CL		: 435/69.1, 2	40.2, 320.1; 53	30/350, 351, 387	7; 536/27.	336/27. Socumentation the Fields Searched Relevant to Claim No. 18 ugiyama nked to 33, see et al., nto two njected entire ollmann member 48, see Masu et otropic 62-763.
II. FIELI	DS SEAR	CHED	Minimum Docu	mentation Searched	4	
Classificati	ion System	•		Classification Symb		
U.S.	. ÷	US CL : 43	5/69.1, 240.2	, 320.1; 530/3	350, 351, 3	187; 536/27.
	٠.				· .	
				d other than Minimur ments are included i	and	
cas, c	online,	aps		,		
						*
III. DOC	UMENTS	CONSIDERED TO B	E RELEVANT 14			- 110
Category*	Citatio	n of Document,16 wit	th indication, where ap	propriate, of the releva	nt passages ¹⁷	Relevant to Claim No. 18
х/ү	et al inosit	., "A new ty	pe of glutama pid metabolis	Pebruary 1987, ate receptor mm, pages 531	linked to	1-3, 6-8/9-30
х/у	"Gluta major	mate recepto: catories: a rat brain mR	r subtypes may study on Xe	1989, Sugiyama be classified nopus oocytes 9-132, see th	into two injected	1-3, 6-8/9-30
У	et al.	. "Cloning b	y functional eceptor famil	ecember 1989, expression of y", pages 643	a member	
x,p	al.,	"sequence a	nd expressio	February 1991, n of a met 765, see pages	abotropic	1-3, 6-30
				m T	¥	
•	. == .	•	·) ·	•		
		•	• • • • •		,	
	•		•	•		
•						·)·
- 40	·		<u> </u>	· · · · · · · · · · · · · · · · · · ·		· .
"A" doct not inter inter inter or vanith doct or vanith doct or	ument defir considered ier docum- mational fili ument which which is cit ther citation ument refer ther means ument publi later than ti	th may throw doubts ted to establish the process of the control of the prior to the interpretation of the priority date claims	of the art which is evance on or after the on priority claim(s) publication date of on (as specified) sure, use, exhibition and od	date or priori application bu theory underly "X" document of invention can considered to "Y" document of invention can invention can invention step one or more of being obvious "&" document me	ty date and no it cited to under ying the invention of the consider involve an inverse particular released to consider involve an inverse particular released to consider the document of the same person skill in the same p	t in conflict with the retend the principle or n evance; the claimed ed novel or cannot be nive step evance; the claimed dered to involve an ment is combined with ents, such combination led in the ert is patent family
Date of th	e Actual C	completion of the Int	ernational Search ²	Date of Mailing of ti	his International	Search Report ²
09	MARCH	1992	· .	\$ 7	33AD 100	9 /
Internation	nal Search	ing Authority ¹		Signature of Author		Vome L
ISF	A/US		· ·	Gian Wang	oh D JIV	1

FURTHER INFORMATION CONTINUED FROM PREVIOUS SHEETS (Not for publication)

I. CLASSIFICATION OF SUBJECT MATTER: IPC (5):

C12P 21/06; C12N 5/00, 15/00; C07H 15/12, 17/00; C07K 3/00, 13/00, 15/00, 17/00; A61K 35/14.

VI. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING This ISA found multiple inventions as follows:

Detailed reasons for holding lack of invention

The claims of the three groups have the characteristics of three distinct inventive concepts. Groups I-III are separate and distinct inventions, and require materially different considerations and searches.

Itemized summary of claims groupings

I. Claims 1-3 and 6-30 are drawn to a method for producing a mammalian G protein by using its encoding sequence, classified in Class 435, subclass 69.1, 240.2; Class 530, subclass 387; Class 536, Subclass 27.

II. Claims 4-5 and 31-33 are drawn to a method for determining the presence of a mammalian G protein by using monoclonal antibody, classified in Class 435, subclass 7,21; Class 424, subclass 85.8.

III. Claims 34-38 are drawn to a method for identifying a compound, classified in Class 435, subclass 4.

Foley & Lardn r Time and Servic s R port

Printed on Q5/18/2001 by Day, Joy A (JADAY)

ate::Thursday, N	ay 17, 2001		
Client: 02606	3 AMERSHAM PHARMACIA BIOTECH	Time:	4.0
Matter: 0101	LEE (USDC CASE NO. C-00-1937) (030516.0062)		
Phase:			
Task: 1	Normal time entry.		• .
Prac Grp: 208	Intellectual Property / Litigation	ē	
	nd finalize papers for filing interference; meetings with R. Warburg to ss strategy.	Open	
		T =:	
Client: 02600		Time:	1.9
Matter: 0102	MD (USDC #C98-01015 & C98-04167) (030516.0006)		× .
Phase:			0
Task: 1	Normal time entry.		. *
Prac Grp: 208	Intellectual Property / Litigation	-	•
	w of Perkin-Elmer documents to determine earliest date Perkin-Elmer knew gabase.	Open	
Client: 0286		Time:	1.1
Matter: 0303	DALHO1130.US.2 -		
Phase:		- 500	:
Task:			.50
Prac Grp: 208	Intellectual Property / Litigation	-	
varrative: Final	editing of declaration of Wright and Pohajdak.	Open	
Client: 99980	00 ZZZ FOLEY & LARDNER-PROFESSIONAL DEVELOPMENT	Time:	2.5
Matter: 0301	INTELLECTUAL PROPERTY DEPARTMENT - GENERAL		
Phase:			
Task:		·	
Prac Grp:		· .	
Narrative: Attend	d SDIPLA meeting.	Open	
		20.000()	7.0
	Totals for Thursday, May 17, 2001: BillableTime(7	•	7.0
*	Non-BillableTime(2		2.5
*	I Ota	al Time:	9.5
	Report Total: BillableTime(7	······································	7.0
	Non-BillableTime(2		2.5
4		U. U. /U/.	